Stochastic ODEs and PDEs for interacting multi-type populations

by

Sandra Martina Kliem

A THESIS SUBMITTED IN PARTIAL FULFILMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

in

The Faculty of Graduate Studies

(Mathematics)

THE UNIVERSITY OF BRITISH COLUMBIA

(Vancouver)

September, 2009

Abstract

This thesis consists of the manuscripts of three research papers studying stochastic ODEs (ordinary differential equations) and PDEs (partial differential equations) that arise in biological models of interacting multi-type populations.

In the first paper I prove uniqueness of the martingale problem for a degenerate SDE (stochastic differential equation) modelling a catalytic branching network. This work is an extension of a paper by Dawson and Perkins to arbitrary networks. The proof is based upon the semigroup perturbation method of Stroock and Varadhan. In the proof estimates on the corresponding semigroup are given in terms of weighted Hölder norms, which are equivalent to a semigroup norm in this generalized setting. An explicit representation of the semigroup is found and estimates using cluster decomposition techniques are derived.

In the second paper I investigate the long-term behaviour of a special class of the SDEs considered above, involving catalytic branching and mutation between types. I analyse the behaviour of the overall sum of masses and the relative distribution of types in the limit using stochastic analysis. For the latter existence, uniqueness and convergence to a stationary distribution are proved by the reasoning of Dawson, Greven, den Hollander, Sun and Swart. One-dimensional diffusion theory allows for a complete analysis of the two-dimensional case.

In the third paper I show that one can construct a sequence of rescaled perturbations of voter processes in $d = 1$ whose approximate densities are tight. This is an extension of the results of Mueller and Tribe for the voter model. We combine critical long-range and fixed kernel interactions in the perturbations. In the long-range case, the approximate densities converge to a continuous density solving a class of SPDEs (stochastic PDEs). For integrable initial conditions, weak uniqueness of the limiting SPDE is shown by a Girsanov theorem. A special case includes a class of stochastic spatial competing species models in mathematical ecology. Tightness is established via a Kolmogorov tightness criterion. Here, estimates on the moments of small increments for the approximate densities are derived via an approximate martingale problem and Green's function representation.

Table of Contents

Appendices

List of Figures

Acknowledgements

I am tremendously grateful to my Ph.D. supervisor Ed Perkins for his help. He introduced me to the problems at hand and provided me with helpful comments and suggestions throughout. Particular thanks go to him for his valuable feedback after reading my manuscript.

I also wish to thank Carl Mueller for a discussion on survival and coexistence questions related to the limiting SPDEs I obtained in my third manuscript.

A special thanks goes to the administration and the computer support of my department that made my life easier in many ways.

Finally, a big thanks goes to all the people I met during my studies at UBC, my parents and my other friends for providing me with help and encouragement throughout my studies.

Chapter 1

Introduction

In the following three Chapters I investigate degenerate stochastic ODEs (ordinary differential equations) and SPDEs (stochastic partial differential equations) that arise in biological models of interacting multi-type populations. In the first two Chapters I investigate the behaviour of the respective masses of a finite number of interacting populations in a non-spatial setting, and in the last paper I study two interacting populations only but add a spatial component. The former will result in the consideration of SDEs (stochastic differential equations), the latter in the consideration of limiting SPDEs. The former models can arise from a network of cooperating branching populations which require the presence of other types (catalysts) to reproduce, while the latter class includes scaling limits of ecological models for two types competing for resources.

When investigating biological models for the evolution of populations over time, the common question to answer is that for survival, extinction and coexistence of types. A well-known model for the evolution of the mass of one type of population is Feller's branching diffusion with parameter γ and linear drift, i.e. the unique solution to the SDE

$$dx_t = bx_t dt + \sqrt{2\gamma x_t} dB_t$$

with constants $b \in \mathbb{R}, \gamma \in \mathbb{R}^+$ and B_t a Brownian motion. Such diffusions can be obtained as the limit of a sequence of rescaled Galton-Watson branching processes at criticality (that is, branching processes with an average number of descendents approaching one). By adding a spatial component, one obtains super-Brownian motion with linear drift instead, which is the unique in law solution $u(t,x)$ to the following SPDE

$$\frac{\partial u}{\partial t} = \Delta u + bu + \sqrt{\gamma u}\dot{W},$$

where $\dot{W} = \dot{W}(t,x)$ is space time white noise. Here, Δu models the spatial motion and dispersion of the population and $\sqrt{\gamma u}\dot{W}$ models the stochastic fluctuations in the population size. $u_t = u(t,\cdot)$ can be interpreted as the continuous spatial density of the population at time t.

For both models, the degeneracy in the fluctuation term and its lack of Lipschitz continuity leads to difficulties in establishing uniqueness. In the above, the additivity properties inherent to both models (for example the sum of two independent (b,γ)-Feller-diffusions starting at x_1 respectively x_2 is a (b,γ)-Feller-diffusion starting at $x_1 + x_2$) can be successfully employed to investigate the long-term behaviour of the above SDE, respectively SPDE. For extensive literature on the above the interested reader is referred to Perkins [14].

As a next step, one can consider an equation for each type of population and introduce interactions between types by interlinking the equations. Thereby one can model competition of types for resources but also mutual help between types. As a result, the analysis of the resulting equations becomes more complicated as additivity properties are not present anymore in this context. In the first and third paper of my thesis I shall therefore employ different perturbation methods to derive results on the systems at hand from results of more accessible models. For instance, the first paper uses a perturbation method of Stroock and Varadhan to obtain the new system of SDEs as a perturbation of a system of independent Feller diffusions with constant coefficients. The last paper considers perturbations of the biased voter model for which the long-range limit was obtained in Mueller and Tribe [12].

In what follows, I shall give a short overview of the models, objectives and underlying literature of each of the three manuscripts of this thesis.

1.1 Overview of the Manuscripts

1.1.1 Degenerate stochastic differential equations for catalytic branching networks

In the first paper I investigate weak uniqueness of solutions to the following system of SDEs: For $j \in R \subset \{1,\dots,d\}$ and $C_j \subset \{1,\dots,d\}\backslash\{j\}$:

$$dx_t^{(j)} = b_j(x_t)dt + \sqrt{2\gamma_j(x_t)\left(\sum_{i\in C_j} x_t^{(i)}\right) x_t^{(j)}}\,dB_t^j \tag{1.1}$$

and for $j \notin R$

$$dx_t^{(j)} = b_j(x_t)dt + \sqrt{2\gamma_j(x_t)x_t^{(j)}}\,dB_t^j. \tag{1.2}$$

Here $x_t \in \mathbb{R}_+^d$ and $b_j, \gamma_j, j = 1,\dots,d$ are Hölder-continuous functions on $\mathbb{R}_+^d$ with $\gamma_j(x) > 0$, and $b_j(x) \geq 0$ if $x_j = 0$. The $B_t^j, j \in \{1,\dots,d\}$ are independent Brownian motions.

This system of SDEs models catalytic branching networks, where types $i \in C_j$ catalyze the replication of type $j, j \in R$ (the so-called reactants). Such systems can be obtained as a limit of near-critical branching particle systems. The growth rate of types corresponds to the branching rate in this stochastic setting, i.e. type $j, j \in R$ in state x_t branches at a rate $\gamma_j(x_t)\sum_{i\in C_j} x_t^{(i)}$ proportional to the sum of masses of types $i, i \in C_j$ at time t.

The degeneracies in the covariance coefficients of this system and their lack of Lipschitz continuity make the investigation of uniqueness a challenging question. The former rules out the classic Stroock-Varadhan approach of perturbing

Brownian motion and the latter prevents application of Itô's pathwise uniqueness arguments. Similar results have been proven in Athreya, Barlow, Bass and Perkins [1] and Bass and Perkins [2] but without the additional singularity $\sum_{i\in C_j} x_t^{(i)}$ in the covariance coefficients of the diffusion.

The question of uniqueness of equations with non-constant coefficients arises already in the case $d = 2$ in the renormalization analysis of hierarchically interacting two-type branching models treated in Dawson, Greven, den Hollander, Sun and Swart [6].

In [7], Dawson and Perkins proved weak uniqueness for the following system of SDEs: For $j \in R$ and $C_j = \{c_j\}, c_j \neq j$,

$$dx_t^{(j)} = b_j(x_t)dt + \sqrt{2\gamma_j(x_t)x_t^{(c_j)}x_t^{(j)}}dB_t^j,$$

and (1.2) as above. This restriction to at most one catalyst per reactant is sufficient for the renormalization analysis for $d = 2$ types, but for more than 2 types one will encounter models where one type may have two catalysts. The goal of my first paper was to overcome this restriction and to allow consideration of general multi-type branching networks as envisioned in the section on future challenges in [6]. In my first paper, I extend the techniques of [7] to the setting of general catalytic networks, i.e. to (1.1) and (1.2). My work further includes natural settings such as competing hypercycles (cf. Eigen and Schuster [8], p.55 respectively Hofbauer and Sigmund [10], p.106). This latter work proposed an analogous system of ODEs as a model for the emergence of long polynucleotides in prebiotic evolution.

1.1.2 Long-term behaviour of a cyclic catalytic branching system

As an application of the above I investigate the following special case in my second paper, involving cyclic catalytic branching and mutation between types. As I shall point out in this paper, the cyclic setup can easily be extended to arbitrary networks. Questions for survival and coexistence of types in the long time limit arise. Such questions naturally arise in biological competition models. For instance, Fleischmann and Xiong [9] investigated a cyclically catalytic super-Brownian motion in one spatial dimension. They showed global segregation (noncoexistence) of neighbouring types in the limit and other results on the finite time survival-extinction but they were not able to determine, if the overall sum dies out in the limit or not. In [10] (p. 86) multi-type branching processes with independent replication and mutation between types were rejected as a model since typically one type would take over, contrary to the observed diversity which emerged from the primordial soup.

Let the following system of SDEs for $d \geq 2$ be given:

$$dx_t^i = \sqrt{2\gamma^i x_t^i x_t^{i+1}}dB_t^i + \sum_{j=1}^{d} x_t^j q_{ji}dt,\ i \in \{1, \ldots, d\},$$

where $x_t^{d+1} \equiv x_t^1$. I assumed the γ^i and $q_{ji}, i \neq j$ to be given positive constants and the $x_0^i \geq 0, i \in \{1, \ldots, d\}$ to be given initial conditions. (q_{ji}) is a Q-matrix modelling the mutations or migrations from type j to type i. I shall investigate in particular the behaviour of the sum of types $s_t = \sum_{i=1}^d x_t^i$ and of the normalized processes $y_t^i = x_t^i / s_t$ in the time-limit. The latter addresses the diversity issue in [10].

1.1.3 Convergence of rescaled competing species processes to a class of SPDEs

The objectives of this paper were threefold. To better understand them, I shall first introduce the three papers that provide motivation.

Firstly, in [12], Mueller and Tribe construct a sequence of rescaled competing species processes $\xi_t^N \in \{0,1\}^{\mathbb{Z}/N}$ in dimension $d = 1$ and show that its approximate densities

$$A(\xi_t^N)(x) \equiv \frac{1}{|\{y \in \mathbb{Z}/N : 0 < |y-x| \leq 1/\sqrt{N}\}|} \sum_{\substack{y \in \mathbb{Z}/N, \\ 0 < |y-x| \leq 1/\sqrt{N}}} \xi_t^N(y), \; x \in N^{-1}\mathbb{Z}$$

converge in distribution to a continuous space time density that solves an SPDE. Here, $\xi_t^N \in \{0,1\}^{\mathbb{Z}/N}$ denotes the configuration at time t of a "voter process" with bias $\tau = \frac{\theta}{N}$. That is, each type (0 or 1) invades a randomly chosen "neighbouring site" with constant rate, where $\theta > 0$ would slightly favour 1's by giving them a slightly larger invasion rate. $\xi_t^N(x) = i$ if site $x \in \mathbb{Z}/N$ is occupied by type $i, i = 0, 1$ and hence u_t can be interpreted as the limiting continuous space time density of type 1 and $1 - u_t$ as the density of type 0. [12] fix $\theta \geq 0$, i.e. consider the case where the opinion of type 1 is slightly dominant. They show that u_t is a solution of the following SPDE, the heat equation with drift, driven by Fisher-Wright noise, namely

$$\frac{\partial u}{\partial t} = \frac{\Delta u}{6} + 2\theta(1-u)u + \sqrt{4u(1-u)}\dot{W}. \tag{1.3}$$

Observe that [12] scale space by $1/N$ and use $\{y \in \mathbb{Z}/N : 0 < |y-x| \leq 1/\sqrt{N}\}$, the set of neighbours of x, to calculate approximate densities. Hence, the number of neighbours of $x \in \mathbb{Z}/N$ is increasing proportionally to $2N^{1/2}$. They thus obtain long-range interactions. Finally, they also speed up time to obtain appropriate limits.

After rescaling (1.3) appropriately, one obtains the Kolmogorov-Petrovskii-Piscuinov (KPP) equation driven by Fisher-Wright noise. The behaviour of this SPDE has already been investigated in Mueller and Sowers [11] in detail, where the existence of travelling waves was shown.

Secondly, in Cox and Perkins [4] it was shown that stochastic spatial Lotka-Volterra models, suitably rescaled in space and time, converge weakly to super-Brownian motion with linear drift. As they choose the parameters in their models to approach 1, the models can also be interpreted as small perturbations of the voter model. [4] extended the main results of Cox, Durrett and Perkins [3], which proved similar results for long-range voter models. Both papers treat the low density regime, i.e. where only a finite number of individuals of type 1 is present. Also note that both papers use a different scaling in comparison to [12]. [12] is at the threshold of the results in [3], but not included, and therefore [12] obtains a non-linear drift term in the limiting SPDE as a result.

[4] considers fixed kernel models in dimensions $d \geq 3$ and long-range kernel models in arbitrary dimension separately. Finally, in Cox and Perkins [5], the results of [4] for $d \geq 3$ are used to relate the limiting super-Brownian motions in the fixed kernel case to questions of coexistence and survival of a rare type in the original Lotka-Volterra model.

Thirdly, spatial versions of the Lotka-Volterra model with finite range were introduced and investigated in Neuhauser and Pacala [13]. The model from [13] incorporates a fecundity parameter and models both intra- and interspecific competition. The paper shows that short-range interactions alter the predictions of the mean-field model.

In my paper I try to extend the approach of [12] for voter models to small perturbations of voter models similar to the perturbations in [4]. I work at criticality in the hope to obtain continuous densities in the limit that solve a class of SPDEs, similar to (1.3) but with more diverse drifts.

My second goal is to thereby include spatial versions of Lotka-Volterra models for competition and fecundity parameters near one as introduced in [13] as the approximating models and to determine their limits. As an additional extension to [12] I shall investigate the weak uniqueness of the limiting class of SPDEs as weak uniqueness of the solutions to the SPDE will yield in turn weak uniqueness of the limits of the approximating densities.

The last objective of this paper was to combine both long-range models at criticality and fixed kernel models in the perturbations. I investigate if the additional fixed kernel perturbation impacts statements on tightness (equivalent to relative compactness in my Polish spaces) of the approximating models. Thereby, results of [4] are extended. As a special case I would then be able to consider rescaled Lotka-Volterra models with long-range dispersal and short-range competition.

1.2 Concluding Remarks

Tightness of approximating particle systems can be used to prove existence of limiting points of the approximating particle systems. Often, all limits can be shown to have certain properties in common. For instance, if all limits satisfy an SDE or SPDE as in my third paper for the case of long-range interactions only, weak uniqueness for the limiting systems of SDEs or SPDEs then yields uniqueness of the limits. Additionally, weak uniqueness of the solutions makes available certain tools that are used to investigate the behaviour of the systems at hand. Therefore, the proof of weak uniqueness in the first paper was fundamental to the analysis of the model for cyclic catalytic branching and mutation between types of the second paper.

All three papers of my thesis have in common that they investigate multi-type interaction models with a degeneracy in the component modelling fluctuations that stems from catalytic branching (the Fisher-Wright noise term can be seen as an application of a 2-cyclic model). Additionally all three models are parameter-dependent, where the parameters can be used to answer questions of survival and coexistence of types.

Bibliography

[1] Athreya, S.R. and Barlow, M.T. and Bass, R.F. and Perkins, E.A. Degenerate stochastic differential equations and super-Markov chains. *Probab. Theory Related Fields* (2002) **123**, 484–520. MR1921011

[2] Bass, R.F. and Perkins, E.A. Degenerate stochastic differential equations with Hölder continuous coefficients and super-Markov chains. *Trans. Amer. Math. Soc.* (2003) **355**, 373–405 (electronic). MR1928092

[3] Cox, J.T. and Durrett, R. and Perkins, E.A. Rescaled voter models converge to super-Brownian motion. *Ann. Probab.* (2000) **28**, 185–234. MR1756003

[4] Cox, J.T. and Perkins, E.A. Rescaled Lotka-Volterra Models converge to Super-Brownian Motion. *Ann. Probab.* (2005) **33**, 904–947. MR2135308

[5] Cox, J.T. and Perkins, E.A. Survival and coexistence in stochastic spatial Lotka-Volterra models. *Probab. Theory Related Fields* (2007) **139**, 89–142. MR2322693

[6] Dawson, D.A. and Greven, A. and den Hollander, F. and Sun, R. and Swart, J.M. The renormalization transformation for two-type branching models. *Ann. Inst. H. Poincaré Probab. Statist.* (2008) **44**, 1038–1077. MR2469334

[7] Dawson, D.A. and Perkins, E.A. On the uniqueness problem for catalytic branching networks and other singular diffusions. *Illinois J. Math.* (2006) **50**, 323–383 (electronic). MR2247832

[8] Eigen, M. and Schuster, P. *The Hypercycle: a principle of natural self-organization.* Springer, Berlin, 1979.

[9] Fleischmann, K. and Xiong, J. A cyclically catalytic super-brownian motion. *Ann. Probab.* (2001) **29**, 820–861. MR1849179

[10] Hofbauer, J. and Sigmund, K. *The Theory of Evolution and Dynamical Systems.* London Math. Soc. Stud. Texts, vol. 7, Cambridge Univ. Press, Cambridge, 1988. MR1071180

[11] Mueller, C. and Sowers, R.B. Random Travelling Waves for the KPP Equation with Noise. *J. Funct. Anal.* (1995) **128**, 439–498. MR1319963

[12] Mueller, C. and Tribe, R. Stochastic p.d.e.'s arising from the long range contact and long range voter processes. *Probab. Theory Related Fields* (1995) **102**, 519–545. MR1346264

[13] Neuhauser, C. and Pacala, S.W. An explicitly spatial version of the Lotka-Volterra model with interspecific competition. *Ann. Appl. Probab.* (1999) **9**, 1226–1259. MR1728561

[14] Perkins, E.A. *Dawson-Watanabe superprocesses and measure-valued diffusions.* Lectures on Probability Theory and Statistics (Saint-Flour, 1999), 125–324, Lecture Notes in Math., 1781, Springer, Berlin, 2002. MR1915445

Chapter 2

Degenerate Stochastic Differential Equations for Catalytic Branching Networks[1]

2.1 Introduction

2.1.1 Catalytic branching networks

In this paper we investigate weak uniqueness of solutions to the following system of stochastic differential equations (SDEs): For $j \in R \subset \{1, \ldots, d\}$ and $C_j \subset \{1, \ldots, d\} \backslash \{j\}$:

$$dx_t^{(j)} = b_j(x_t)dt + \sqrt{2\gamma_j(x_t)\left(\sum_{i \in C_j} x_t^{(i)}\right) x_t^{(j)}} dB_t^j \tag{2.1}$$

and for $j \notin R$

$$dx_t^{(j)} = b_j(x_t)dt + \sqrt{2\gamma_j(x_t)x_t^{(j)}} dB_t^j. \tag{2.2}$$

Here $x_t \in \mathbb{R}_+^d$ and $b_j, \gamma_j, j = 1, \ldots, d$ are Hölder-continuous functions on $\mathbb{R}_+^d$ with $\gamma_j(x) > 0$, and $b_j(x) \geq 0$ if $x_j = 0$.

The degeneracies in the covariance coefficients of this system make the investigation of uniqueness a challenging question. Similar results have been proven in [1] and [4] but without the additional singularity $\sum_{i \in C_j} x_t^{(i)}$ in the covariance coefficients of the diffusion. Other types of singularities, for instance replacing the additive form by a multiplicative form $\prod_{i \in C_j} x_t^{(i)}$, are possible as well, under additional assumptions on the structure of the network (cf. Remark 2.1.9 at the end of Subsection 2.1.5).

The given system of SDEs can be understood as a stochastic analogue to a system of ODEs for the concentrations $y_j, j = 1, \ldots, d$ of a type T_j. Then

[1] A version of this chapter has been accepted for publication. Kliem, S.M. (2009) Degenerate Stochastic Differential Equations for Catalytic Branching Networks. Ann. Inst. H. Poincaré Probab. Statist.

$y_j/\dot{y}_j$ corresponds to the rate of growth of type T_j and one obtains the following ODEs (see [9]): for independent replication $\dot{y}_j = b_j y_j$, autocatalytic replication $\dot{y}_j = \gamma_j y_j^2$ and catalytic replication $\dot{y}_j = \gamma_j \left(\sum_{i\in C_j} y_i\right) y_j$. In the catalytic case the types $T_i, i \in C_j$ catalyze the replication of type j, i.e. the growth of type j is proportional to the sum of masses of types $i, i \in C_j$ present at time t.

An important case of the above system of ODEs is the so-called hypercycle, firstly introduced by Eigen and Schuster (see [8]). It models hypercyclic replication, i.e. $\dot{y}_j = \gamma_j y_{j-1} y_j$ and represents the simplest form of mutual help between different types.

The system of SDEs can be obtained as a limit of branching particle systems. The growth rate of types in the ODE setting now corresponds to the branching rate in the stochastic setting, i.e. type j branches at a rate proportional to the sum of masses of types $i, i \in C_j$ at time t.

The question of uniqueness of equations with non-constant coefficients arises already in the case $d = 2$ in the renormalization analysis of hierarchically interacting two-type branching models treated in [6]. The consideration of successive block averages leads to a renormalization transformation on the diffusion functions of the SDE

$$dx_t^{(i)} = c\left(\theta_i - x_t^{(i)}\right)dt + \sqrt{2g_i(x_t)}dB_t^i, i = 1, 2$$

with $\theta_i \geq 0, i = 1, 2$ fixed. Here $g = (g_1, g_2)$ with $g_i(x) = x_i \gamma_i(x)$ or $g_i(x) = x_1 x_2 \gamma_i(x)$, $i = 1, 2$ for some positive continuous function γ_i on $\mathbb{R}_+^2$. The renormalization transformation acts on the diffusion coefficients g and produces a new set of diffusion coefficients for the next order block averages. To be able to iterate the renormalization transformation indefinitely a subclass of diffusion functions has to be found that is closed under the renormalization transformation. To even define the renormalization transformation one needs to show that the above SDE has a unique weak solution and to iterate it we need to establish uniqueness under minimal conditions on the coefficients.

This paper is an extension of the work done in Dawson and Perkins [7]. The latter, motivated by the stochastic analogue to the hypercycle and by [6], proved weak uniqueness in the above mentioned system of SDEs (2.1) and (2.2), where (2.1) is restricted to

$$dx_t^{(j)} = b_j(x_t)dt + \sqrt{2\gamma_j(x_t)x_t^{(c_j)}x_t^{(j)}}dB_t^j,$$

i.e. $C_j = \{c_j\}$ and (2.2) remains unchanged. This restriction to at most one catalyst per reactant is sufficient for the renormalization analysis for $d = 2$ types, but for more than 2 types one will encounter models where one type may have two catalysts. The present work overcomes this restriction and allows consideration of general multi-type branching networks as envisioned in [6], including further natural settings such as competing hypercycles (cf. [8] page 55 resp. [9], p. 106). In particular, the techniques of [7] will be extended to the setting of general catalytic networks.

Intuitively it is reasonable to conjecture uniqueness in the general setting as there is less degeneracy in the diffusion coefficients; $x_t^{(c_j)}$ changes to $\sum_{i\in C_j} x_t^{(i)}$,

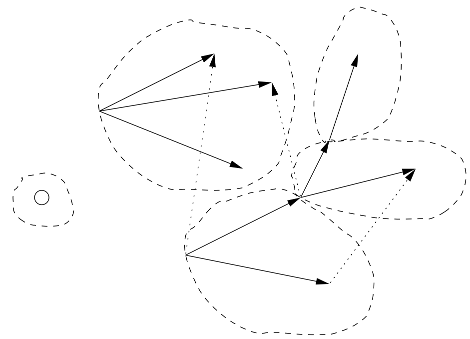

Figure 2.1: Decomposition from the catalyst's point of view: Arrows point from vertices $i \in N_C$ to vertices $j \in R_i$ (for the definition of N_C, R_i and N_2 see Subsections 2.1.3 and 2.1.5 to follow). Separate points signify vertices $j \in N_2$. The dotted arrows signify arrows which are only allowed in the generalized setting and thus make a decomposition of the kind used in [7] inaccessible.

all coordinates $i \in C_j$ have to become zero at the same time to result in a singularity.

For $d = 2$ weak uniqueness was proven for a special case of a mutually catalytic model ($\gamma_1 = \gamma_2 = \text{const.}$) via a duality argument in [10]. Unfortunately this argument does not extend to the case $d > 2$.

2.1.2 Comparison with Dawson and Perkins [7]

The generalization to arbitrary networks results in more involved calculations. The most significant change is the additional dependency among catalysts. In [7] the semigroup of the process under consideration could be decomposed into groups of single vertices and groups of catalysts with their corresponding reactants (see Figure 2.1). Hence the main part of the calculations in [7], where bounds on the semigroup are derived, i.e. Section 2 of [7] ("Properties of the basic semigroups"), could be reduced to the setting of a single vertex or a single catalyst with a finite number of reactants. In the general setting this strategy is no longer available as one reactant is now allowed to have multiple catalysts (see again Figure 2.1). As a consequence we shall treat all vertices in one step only. This results in more work in Section 2, where bounds on the given semigroup are now derived directly.

We also employ a change of perspective from reactants to catalysts. In [7] every reactant j had one catalyst c_j only (and every catalyst i a set of reactants R_i). For the general setting it turns out to be more efficient to consider every catalyst i with the set R_i of its reactants. In particular, the restriction from R_i to $\bar{R}_i$, including only reactants whose catalysts are all zero, turns out to be crucial for later definitions and calculations. It plays a key role in the extension of the definition of the weighted Hölder norms to general networks (see Subsection 2.1.6).

Changes in one catalyst indirectly impact other catalysts now via common reactants, resulting for instance in new mixed partial derivatives. As a first step a representation for the semigroup of the generalized process had to be found (see (2.15)). In [7], (12) the semigroup could be rewritten in a product form of semigroups of each catalyst with its reactants. Now a change in one catalyst resp. coordinate of the semigroup impacts in particular the local covariance of all its reactants. As the other catalysts of this reactant also appear in this coefficient, a decomposition becomes impossible. Instead the triangle inequality has to be often used to express resulting multi-dimensional coordinate changes of the function G, which is closely related with the semigroup representation (see (2.16)), via one-dimensional ones. As another important tool Lemma 2.2.6 was developed in this context.

The ideas of the proofs in [7] often had to be extended. Major changes can be found in the critical Proposition 2.2.25 and its associated Lemmas (especially Lemma 2.2.29). The careful extension of the weighted Hölder norms to arbitrary networks had direct impact on the proofs of Lemma 2.2.19 and Theorem 2.2.20.

2.1.3 The model

Let a branching network be given by a directed graph $(V, \mathcal{E})$ with vertices $V = \{1, \ldots, d\}$ and a set of directed edges $\mathcal{E} = \{e_1, \ldots, e_k\}$. The vertices represent the different types, whose growth is under investigation, and $(i, j) \in \mathcal{E}$ means that type i "catalyzes" the branching of type j. As in [7] we continue to assume:

Hypothesis 2.1.1. *$(i, i) \notin \mathcal{E}$ for all $i \in V$.*

Let C denote the set of catalysts, i.e. the set of vertices which appear as the 1st element of an edge and R denote the set of reactants, i.e. the set of vertices that appear as the 2nd element of an edge.

For $j \in R$, let

$$C_j = \{i : (i, j) \in \mathcal{E}\}$$

be the set of catalysts of j and for $i \in C$, let

$$R_i = \{j : (i, j) \in \mathcal{E}\}$$

be the set of reactants, catalyzed by i. If $j \notin R$ let $C_j = \emptyset$ and if $i \notin C$, let $R_i = \emptyset$.

We shall consider the following system of SDEs:

For $j \in R$:

$$dx_t^{(j)} = b_j(x_t)dt + \sqrt{2\gamma_j(x_t) \left(\sum_{i \in C_j} x_t^{(i)} \right) x_t^{(j)}} dB_t^j$$

and for $j \notin R$

$$dx_t^{(j)} = b_j(x_t)dt + \sqrt{2\gamma_j(x_t) x_t^{(j)}} dB_t^j.$$

Our goal will be to show the weak uniqueness of the given system of SDEs.

2.1.4 Statement of the main result

In what follows we shall impose additional regularity conditions on the coefficients of our diffusions, similar to the ones in Hypothesis 2 of [7], which will remain valid unless indicated to the contrary. $|x|$ is the Euclidean length of $x \in \mathbb{R}^d$ and for $i \in V$ let e_i denote the unit vector in the ith direction.

Hypothesis 2.1.2. *For $i \in V$,*

$$\gamma_i : \mathbb{R}_+^d \to (0, \infty),$$
$$b_i : \mathbb{R}_+^d \to \mathbb{R}$$

are taken to be Hölder continuous of some positive index on compact subsets of $\mathbb{R}_+^d$ such that $|b_i(x)| \le c(1 + |x|)$ on $\mathbb{R}_+^d$, and

$$\begin{cases} b_i(x) \ge 0 \text{ if } x_i = 0. \text{ In addition,} \\ b_i(x) > 0 \text{ if } i \in C \cup R \text{ and } x_i = 0. \end{cases}$$

Definition 2.1.3. If ν is a probability on $\mathbb{R}_+^d$, a probability P on $\mathcal{C}(\mathbb{R}_+, \mathbb{R}_+^d)$ is said to solve the martingale problem $MP(\mathcal{A},\nu)$ if under P, the law of $x_0(\omega) = \omega_0$ ($x_t(\omega) = \omega(t)$) is ν and for all $f \in \mathcal{C}_b^2(\mathbb{R}_+^d)$,

$$M_f(t) = f(x_t) - f(x_0) - \int_0^t \mathcal{A}f(x_s)ds$$

is a local martingale under P with respect to the canonical right-continuous filtration $(\mathcal{F}_t)$.

Remark 2.1.4. *The weak uniqueness of a system of SDEs is equivalent to the uniqueness of the corresponding martingale problem (see for instance, [12], V.(19.7)).*

For $f \in \mathcal{C}_b^2(\mathbb{R}_+^d)$, the generator corresponding to our system of SDEs is

$$\begin{aligned} \mathcal{A}f(x) &= \mathcal{A}^{(b,\gamma)}f(x) \\ &= \sum_{j \in R} \gamma_j(x) \left(\sum_{i \in C_j} x_i \right) x_j f_{jj}(x) + \sum_{j \notin R} \gamma_j(x) x_j f_{jj}(x) + \sum_{j \in V} b_j(x) f_j(x). \end{aligned}$$

Here f_{ij} is the second partial derivative of f w.r.t. x_i and x_j.

As a state space for the generator $\mathcal{A}$ we shall use

$$\mathcal{S} = \left\{ x \in \mathbb{R}_+^d : \prod_{j \in R} \left(\sum_{i \in C_j} x_i + x_j \right) > 0 \right\}. \tag{2.3}$$

We first note that $\mathcal{S}$ is a natural state space for $\mathcal{A}$.

Lemma 2.1.5. *If P is a solution to MP($\mathcal{A}$,ν), where ν is a probability on $\mathbb{R}^d_+$, then $x_t \in \mathcal{S}$ for all $t > 0$ P-a.s.*

Proof. The proof follows as for Lemma 5, [7] on p. 377 via a comparison argument with a Bessel process, using Hypothesis 2.1.2. □

We shall now state the main theorem which, together with Remark 2.1.4 provides weak uniqueness of the given system of SDEs for a branching network.

Theorem 2.1.6. *Assume Hypothesis 2.1.1 and 2.1.2 hold. Then for any probability ν, on $\mathcal{S}$, there is exactly one solution to MP($\mathcal{A}$,ν).*

2.1.5 Outline of the proof

Our main task in proving Theorem 2.1.6 consists in establishing uniqueness of solutions to the martingale problem *MP*($\mathcal{A}$,ν). Existence can be proven as in Theorem 1.1 of [1]. The main idea in proving uniqueness consists in understanding our diffusion as a perturbation of a well-behaved diffusion and applying the Stroock-Varadhan perturbation method (refer to [13]) to it. This approach can be devided into three steps.

Step 1: Reduction of the problem. We can assume w.l.o.g. that $\nu = \delta_{x^0}$. Furthermore it is enough to consider uniqueness for families of strong Markov solutions. Indeed, the first reduction follows by a standard conditioning argument (see p. 136 of [3]) and the second reduction follows by using Krylov's Markov selection theorem (Theorem 12.2.4 of [13]) together with the proof of Proposition 2.1 in [1].

Next we shall use a localization argument of [13] (see e.g. the argument in the proof of Theorem 1.2 of [4]), which basically states that it is enough if for each $x^0 \in \mathcal{S}$ the martingale problem $MP(\tilde{\mathcal{A}}, \delta_{x^0})$ has a unique solution, where $b_i = \tilde{b}_i$ and $\gamma_i = \tilde{\gamma}_i$ agree on some $B(x^0, r_0) \cap \mathbb{R}^d_+$. Here we used in particular that a solution never exits $\mathcal{S}$ as shown in Lemma 2.1.5.

Finally, if the covariance matrix of the diffusion is non-degenerate, uniqueness follows by a perturbation argument as in [13] (use e.g. Theorem 6.6.1 and Theorem 7.2.1). Hence consider only singular initial points, i.e. where either

$$\left\{x_0^{(j)} = 0 \text{ or } \sum_{i \in C_j} x_0^{(i)} = 0 \text{ for some } j \in R\right\} \text{ or } \left\{x_0^{(j)} = 0 \text{ for some } j \notin R.\right\}$$

Step 2: Perturbation of the generator. Fix a singular initial point $x^0 \in \mathcal{S}$ and set (for an example see Figure 2.2)

$$\begin{aligned} N_R &= \left\{ j \in R : \sum_{i \in C_j} x_i^0 = 0 \right\}; \\ N_C &= \cup_{j \in N_R} C_j; \\ N_2 &= V \setminus (N_R \cup N_C); \\ \bar{R}_i &= R_i \cap N_R, \end{aligned}$$

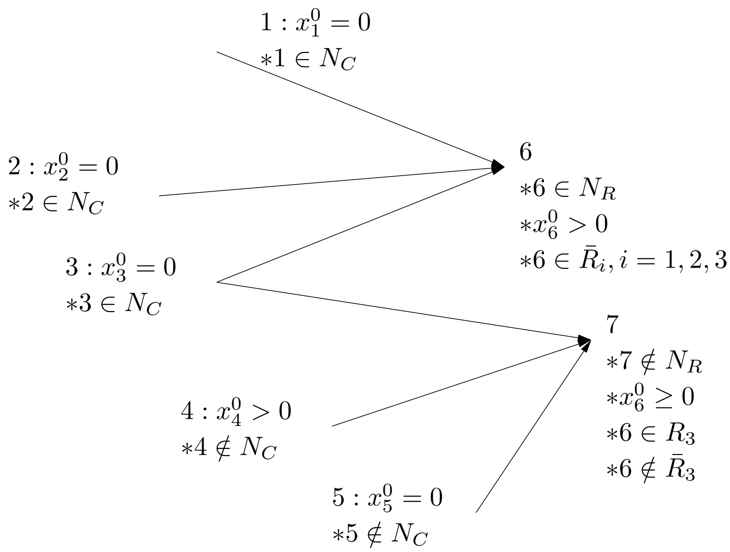

Figure 2.2: Definition of N_R, N_C and $\bar{R}_i$. The $*$'s are the implications deduced from the given setting.

i.e. in contrast to the setting in [7], p. 327, N_2 can also include zero catalysts, but only those whose reactants have at least one more catalyst being non-zero.

Let $Z = Z(x^0) = \{i \in V : x_i^0 = 0\}$ (if $i \notin Z$, then $x_i^0 > 0$ and so $x_s^{(i)} > 0$ for small s a.s. by continuity). Moreover, if $x^0 \in \mathcal{S}$, then $N_R \cap Z = \emptyset$ and

$$N_R \cup N_C \cup N_2 = V$$

is a disjoint union.

Notation 2.1.7. In what follows let

$$\mathbb{R}^A \equiv \{f, f : A \to \mathbb{R}\} \text{ resp. } \mathbb{R}_+^A \equiv \{f, f : A \to \mathbb{R}_+\}.$$

for arbitrary $A \subset V$.

Next we shall rewrite our system of SDEs with corresponding generator $\mathcal{A}$ as a perturbation of a well-understood system of SDEs with corresponding generator $\mathcal{A}^0$, which has a unique solution. The state space of $\mathcal{A}^0$ will be $\mathcal{S}(x^0) = \mathcal{S}_0 = \{x \in \mathbb{R}^d : x_i \geq 0 \text{ for all } i \notin N_R\}$.

First, we view $\left(\left\{x^{(j)}\right\}_{j \in N_R}, \left\{x^{(i)}\right\}_{i \in N_C}\right)$, i.e. the set of vertices with zero catalysts together with these catalysts, near its initial point $\left(\left\{x_j^0\right\}_{j \in N_R}, \left\{x_i^0\right\}_{i \in N_C}\right)$ as a perturbation of the diffusion on $\mathbb{R}^{N_R} \times \mathbb{R}_+^{N_C}$, which is given by the unique solution to the following system of SDEs:

$$dx_t^{(j)} = b_j^0 dt + \sqrt{2\gamma_j^0 \left(\sum_{i \in C_j} x_t^{(i)}\right)} dB_t^j,\ x_0^{(j)} = x_j^0,\ \text{for } j \in N_R$$

and (2.4)

$$dx_t^{(i)} = b_i^0 dt + \sqrt{2\gamma_i^0 x_t^{(i)}} dB_t^i,\ x_0^{(i)} = x_i^0,\ \text{for } i \in N_C,$$

where for $j \in N_R$, $b_j^0 = b_j(x^0) \in \mathbb{R}$ and $\gamma_j^0 = \gamma_j(x^0)x_j^0 > 0$ as $x_j^0 > 0$ if its catalysts are all zero. Also, $b_i^0 = b_i(x^0) > 0$ as $x_i^0 = 0$ for $i \in N_C$ and $\gamma_i^0 = \gamma_i(x^0)\sum_{k \in C_i} x_k^0 > 0$ if $i \in N_C \cap R$ as i is a zero catalyst thus having at least one non-zero catalyst itself, or $\gamma_i^0 = \gamma_i(x^0) > 0$ if $i \in N_C \backslash R$. Note that the non-negativity of $b_i^0, i \in N_C$ ensures that solutions starting in $\{x_i^0 \geq 0\}$ remain there (also see definition of $\mathcal{S}_0$).

Secondly, for $j \in N_2$ we view this coordinate as a perturbation of the Feller branching process (with immigration)

$$dx_t^{(j)} = b_j^0 dt + \sqrt{2\gamma_j^0 x_t^{(j)}} dB_t^j,\ x_0^{(j)} = x_j^0,\ \text{for } j \in N_2, \tag{2.5}$$

where $b_j^0 = (b_j(x^0) \vee 0)$ (at the end of Section 2.3 the general case $b_j(x^0) \in \mathbb{R}$ is reduced to $b_j(x^0) \geq 0$ by a Girsanov transformation), $\gamma_j^0 = \gamma_j(x^0)\sum_{i \in C_j} x_i^0 > 0$ if $j \in R$ by definition of N_2, i.e. at least one of the catalysts being positive, or $\gamma_j^0 = \gamma_j(x^0) > 0$ if $j \notin R$. As for $i \in N_C$, the non-negativity of $b_j^0, j \in N_2$ ensures that solutions starting in $\{x_j^0 \geq 0\}$ remain there (see again definition of $\mathcal{S}_0$).

Therefore we can view $\mathcal{A}$ as a perturbation of the generator

$$\mathcal{A}^0 = \sum_{j \in V} b_j^0 \frac{\partial}{\partial x_j} + \sum_{j \in N_R} \gamma_j^0 \left(\sum_{i \in C_j} x_i\right) \frac{\partial^2}{\partial x_j^2} + \sum_{i \in N_C \cup N_2} \gamma_i^0 x_i \frac{\partial^2}{\partial x_i^2}. \tag{2.6}$$

The coefficients b_i^0, γ_i^0 found above for $x^0 \in \mathcal{S}$ now satisfy

$$\begin{cases} \gamma_j^0 > 0 \text{ for all } j, \\ b_j^0 \geq 0 \text{ if } j \notin N_R, \\ b_j^0 > 0 \text{ if } j \in (R \cup C) \cap Z, \end{cases} \tag{2.7}$$

where

$$N_R \cap Z = \emptyset. \tag{2.8}$$

In the remainder of the paper we shall always assume the conditions (2.7) hold when dealing with $\mathcal{A}^0$ whether or not it arises from a particular $x^0 \in \mathcal{S}$ as above. As we shall see in Subsection 2.2.1 the $\mathcal{A}^0$ martingale problem is then well-posed and the solution is a diffusion on

$$\mathcal{S}_0 \equiv \mathcal{S}(x^0) = \{x \in \mathbb{R}^d : x_i \geq 0 \text{ for all } i \in V \backslash N_R = N_C \cup N_2\}. \tag{2.9}$$

Notation 2.1.8. In the following we shall use the notation

$$N_{C2} \equiv N_C \cup N_2.$$

Step 3: A key estimate. Set

$$\begin{aligned}\mathcal{B}f &:= (\mathcal{A} - \mathcal{A}^0)f \\ &= \sum_{j\in V} \left(\tilde{b}_j(x) - b_j^0\right) \frac{\partial f}{\partial x_j} + \sum_{j\in N_R} \left(\tilde{\gamma}_j(x) - \gamma_j^0\right) \left(\sum_{i\in C_j} x_i\right) \frac{\partial^2 f}{\partial x_j^2} \\ &\quad + \sum_{i\in N_{C2}} \left(\tilde{\gamma}_i(x) - \gamma_i^0\right) x_i \frac{\partial^2 f}{\partial x_i^2},\end{aligned}$$

where

$$\begin{aligned}&\text{for } j \in V, && \tilde{b}_j(x) = b_j(x), \\ &\text{for } j \in N_R, && \tilde{\gamma}_j(x) = \gamma_j(x) x_j, \text{ and} \\ &\text{for } i \in N_{C2}, && \tilde{\gamma}_i(x) = 1_{\{i\in R\}}\gamma_i(x) \sum_{k\in C_i} x_k + 1_{\{i\notin R\}}\gamma_i(x).\end{aligned}$$

By using the continuity of the diffusion coefficients of $\mathcal{A}$ and the localization argument mentioned in Step 1 we may assume that the coefficients of the operator $\mathcal{B}$ are arbitrarily small, say less than η in absolute value. The key step (see Theorem 2.3.3) will be to find a Banach space of continuous functions with norm $\|\cdot\|$, depending on x^0, so that for η small enough and $\lambda_0 > 0$ large enough,

$$\| \mathcal{B}R_\lambda f \| \leq \frac{1}{2} \| f \|, \ \forall\ \lambda > \lambda_0. \tag{2.10}$$

Here

$$R_\lambda f = \int_0^\infty e^{-\lambda s} P_s f ds \tag{2.11}$$

is the resolvent of the diffusion with generator $\mathcal{A}^0$ and P_t is its semigroup.

The uniqueness of the resolvent of our strong Markov solution will then follow as in [13] and [4]. A sketch of the proof is given in Section 2.3.

Remark 2.1.9. *Under additional restrictions on the structure of the branching network our results carry over to the system of SDEs, where the additive form for the catalysts is replaced by a multiplicative form as follows. For $j \in R$ we now consider*

$$dx_t^{(j)} = b_j(x_t)dt + \sqrt{2\gamma_j(x_t)\left(\prod_{i\in C_j} x_t^{(i)}\right) x_t^{(j)}} dB_t^j$$

instead and for $j \notin R$

$$dx_t^{(j)} = b_j(x_t)dt + \sqrt{2\gamma_j(x_t)x_t^{(j)}} dB_t^j$$

as before. Indeed, if we impose that for all $j \in R$ we have either

$$|C_j| = 1 \text{ or}$$
$$|C_j| \geq 2 \text{ and for all } i_1 \neq i_2, i_1, i_2 \in C_j : i_1 \in C_{i_2} \text{ or } i_2 \in C_{i_1},$$

and if we assume that Hypothesis 2.1.2 holds, then we can show a result similar to Theorem 2.1.6.

For instance, the following system of SDEs would be included.

$$\begin{aligned} dx_t^{(1)} &= b_1(x_t)dt + \sqrt{2\gamma_1(x_t)x_t^{(2)}x_t^{(3)}x_t^{(1)}}dB_t^1, \\ dx_t^{(2)} &= b_2(x_t)dt + \sqrt{2\gamma_2(x_t)x_t^{(3)}x_t^{(4)}x_t^{(2)}}dB_t^2, \\ dx_t^{(3)} &= b_3(x_t)dt + \sqrt{2\gamma_3(x_t)x_t^{(4)}x_t^{(1)}x_t^{(3)}}dB_t^3, \\ dx_t^{(4)} &= b_4(x_t)dt + \sqrt{2\gamma_4(x_t)x_t^{(1)}x_t^{(2)}x_t^{(4)}}dB_t^4. \end{aligned}$$

Note in particular, that the additional assumptions on the network ensure that at most one of either the catalysts in C_j or j itself can become zero, so that we obtain the same generator $\mathcal{A}^0$ as in the setting of additive catalysts if we set $\gamma_j^0 \equiv \gamma_j(x^0) \prod_{i \in \{j\} \cup C_j : x_i^0 > 0} x_i^0$ (cf. the derivation of (2.4)).

Remark 2.1.10. *In [5] the Hölder condition on the coefficients was successfully removed but the restrictions on the network as stated in [7] were kept. As both [7] and [5] are based upon realizing the SDE in question as a perturbation of a well-understood SDE, one could start extending [5] to arbitrary networks by using the same generator and semigroup decomposition for the well-understood SDE as considered in this paper.*

2.1.6 Weighted Hölder norms and semigroup norms

In this section we describe the Banach space of functions which will be used in (2.10). In (2.10) we use the resolvent of the generator $\mathcal{A}^0$ with state space $\mathcal{S}_0 = \mathcal{S}(x^0) = \{x \in \mathbb{R}^d : x_i \geq 0 \text{ for all } i \in N_{C2}\}$. Note in particular that the state space and the realizations of the sets $N_R, \bar{R}_i$ etc. depend on x^0.

Next we shall define the *Banach space of weighted α-Hölder continuous functions on $\mathcal{S}_0$, $\mathcal{C}_w^\alpha(\mathcal{S}_0) \subset \mathcal{C}_b(\mathcal{S}_0)$*, in two steps. It will be the Banach space we look for and is a modification of the space of weighted Hölder norms used in [4].

Let $f : \mathcal{S}_0 \to \mathbb{R}$ be bounded and measurable and $\alpha \in (0,1)$. As a first step define the following seminorms for $i \in N_C$:

$$\begin{aligned} |f|_{\alpha,i} = \sup\Big\{ &|f(x+h) - f(x)| \left(|h|^{-\alpha} x_i^{\alpha/2} \vee |h|^{-\alpha/2} \right) : \\ &|h| > 0, h_k = 0 \text{ if } k \notin \{i\} \cup \bar{R}_i, x, h \in \mathcal{S}_0 \Big\}. \end{aligned}$$

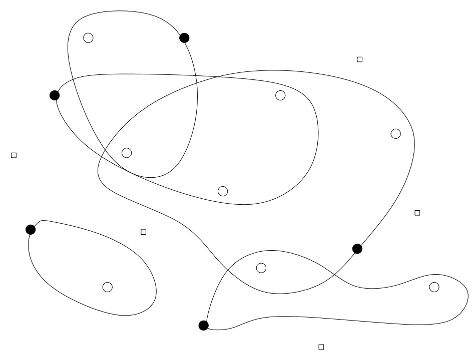

Figure 2.3: Decomposition of the system of SDEs: unfilled circles, resp. filled circles, resp. squares are elements of N_R, resp. N_C, resp. N_2. The definition of $|f|_{\alpha,i}, i \in N_C$ allows changes in i (filled circles) and the associated $j \in \bar{R}_i$ (unfilled circles), the definition of $|f|_{\alpha,j}, j \in N_2$ allows changes in $j \in N_2$ (squares). Hence changes in all vertices are possible.

For $j \in N_2$ this corresponds to setting

$$|f|_{\alpha,j} = \sup\Big\{ |f(x+h) - f(x)| \left(|h|^{-\alpha} x_j^{\alpha/2} \vee |h|^{-\alpha/2} \right) : \\ h_j > 0, h_k = 0 \text{ if } k \neq j, x \in \mathcal{S}_0 \Big\}.$$

This definition is an extension of the definition in [7], p. 329. In our context the definition of $|f|_{\alpha,i}, i \in N_C$ had to be extended carefully by replacing the set R_i (in [7] equal to the set $\bar{R}_i$) by the set $\bar{R}_i \subset R_i$. Observe that the seminorms for $i \in N_C$ and $j \in N_2$ taken together still allow changes in all coordinates (see Figure 2.3). The definition of $|f|_{\alpha,j}, j \in N_2$ furthermore varies slightly from the one in [7]. We use our definition instead as it enables us to handle the coordinates $i \in N_C, j \in N_2$ without distinction.

Secondly, set $I = N_{C2}$. Then let

$$|f|_{\mathcal{C}_w^\alpha} = \max_{j \in I} |f|_{\alpha,j}, \qquad \| f \|_{\mathcal{C}_w^\alpha} = |f|_{\mathcal{C}_w^\alpha} + \| f \|_\infty,$$

where $\| f \|_\infty$ is the supremum norm of f. $\| f \|_{\mathcal{C}_w^\alpha}$ is the norm we looked for and its corresponding Banach subspace of $\mathcal{C}_b(\mathcal{S}_0)$ is

$$\mathcal{C}_w^\alpha(\mathcal{S}_0) = \{ f \in \mathcal{C}_b(\mathcal{S}_0) : \| f \|_{\mathcal{C}_w^\alpha} < \infty \},$$

the Banach space of weighted α-Hölder continuous functions on $\mathcal{S}_0$. Note that the definition of the seminorms $|f|_{\alpha,j}, j \in I$ depends on $N_C, \bar{R}_i$ etc. and hence on x^0. Thus $\| f \|_{\mathcal{C}_w^\alpha}$ depends on x^0 as well.

The seminorms $|f|_{\alpha,i}$ are weaker norms near the spatial degeneracy at $x_i = 0$ where we expect to have less smoothing by the resolvent.

Some more background on the choice of the above norms can be found in [4], Section 2. Bass and Perkins ([4]) consider

$$|f|^*_{\alpha,i} \equiv \sup\Big\{|f(x+he_i)-f(x)||h|^{-\alpha}x_i^{\alpha/2} : h>0, x\in\mathbb{R}^d_+\Big\},$$
$$|f|^*_\alpha \equiv \sup_{i\le d}|f|^*_{\alpha,i} \text{ and } \|f\|^*_\alpha \equiv |f|^*_\alpha + \|f\|_\infty$$

instead, where e_i denotes the unit vector in the i-th direction in $\mathbb{R}^d$. They show that if $f\in\mathcal{C}_b(\mathbb{R}^d_+)$ is uniformly Hölder of index $\alpha\in(0,1]$, and constant outside of a bounded set, then $f\in\mathcal{C}^{\alpha,*}_w \equiv \{f\in\mathcal{C}_b(\mathbb{R}^d_+) : \|f\|^*_\alpha<\infty\}$. On the other hand, $f\in\mathcal{C}^{\alpha,*}_w$ implies f is uniformly Hölder of order $\alpha/2$.

As it will turn out later (see Theorem 2.2.20) our norm $\|f\|_{\mathcal{C}^\alpha_w}$ is equivalent to another norm, the so-called *semigroup norm*, defined via the semigroup P_t corresponding to the generator $\mathcal{A}^0$ of our process. As we shall mainly investigate properties of the semigroup P_t on $\mathcal{C}_b(\mathcal{S}_0)$ in what follows, it is not surprising that this equivalence turns out to be useful in later calculations.

In general one defines the semigroup norm (cf. [2]) for a Markov semigroup $\{P_t\}$ on the bounded Borel functions on D where $D\subset\mathbb{R}^d$ and $\alpha\in(0,1)$ via

$$|f|_\alpha = \sup_{t>0}\frac{\|P_tf-f\|_\infty}{t^{\alpha/2}}, \qquad \|f\|_\alpha = |f|_\alpha + \|f\|_\infty\,. \tag{2.12}$$

The associated Banach space of functions is then

$$\mathcal{S}^\alpha = \{f : D\to\mathbb{R} : f \text{ Borel }, \|f\|_\alpha<\infty\}. \tag{2.13}$$

Convention 2.1.11. *Throughout this paper all constants appearing in statements of results and their proofs may depend on a fixed parameter $\alpha\in(0,1)$ and $\{b^0_j,\gamma^0_j : j\in V\}$ as well as on $|V|=d$. By (2.7)*

$$M^0 = M^0(\gamma^0,b^0) \equiv \max_{i\in V}\Big\{\gamma^0_i \vee \big(\gamma^0_i\big)^{-1} \vee \big|b^0_i\big|\Big\} \vee \max_{i\in(R\cup C)\cap Z}\big(b^0_i\big)^{-1} < \infty. \tag{2.14}$$

Given $\alpha\in(0,1)$, d and $0<M<\infty$, we can, and shall, choose the constants to hold uniformly for all coefficients satisfying $M^0\le M$.

2.1.7 Outline of the paper

Proofs only requiring minor adaptations from those in [7] are usually omitted. A more extensive version of the proofs appearing in Sections 2.2 and 2.3 may be found on the arXiv at arXiv:0802.0035v2.

The outline of the paper is as follows. In Section 2.2 the semigroup P_t corresponding to the generator $\mathcal{A}^0$ on the state space $\mathcal{S}_0$, as introduced in (2.6) and (2.9), will be investigated. We start with giving an explicit representation of the semigroup in Subsection 2.2.1. In Subsection 2.2.2 the canonical measure $\mathbb{N}_0$ is introduced which is used in Subsection 2.2.3 to prove existence and give a representation of derivatives of the semigroup. In Subsections 2.2.4 and 2.2.5

bounds are derived on the L^∞ norms and on the weighted Hölder norms of those differentiation operators applied to $P_t f$, which appear in the definition of $\mathcal{A}^0$. Furthermore, at the end of Subsection 2.2.4 the equivalence of the weighted Hölder norm and semigroup norm is shown. Finally, in Section 2.3 bounds on the resolvent R_λ of P_t are deduced from the bounds on P_t found in Section 2.2. The bounds on the resolvent will then be used to obtain the key estimate (2.10). The remainder of Section 2.3 illustrates how to prove the uniqueness of solutions to the martingale problem $MP(\mathcal{A},\nu)$ from this, as in [7].

2.2 Properties of the Semigroup

2.2.1 Representation of the semigroup

In this subsection we shall find an explicit representation of the semigroup P_t corresponding to the generator $\mathcal{A}^0$ (cf. (2.6)) on the state space $\mathcal{S}_0$ and further preliminary results. We assume the coefficients satisfy (2.7) and Convention 2.1.11 holds.

Let us have a look at (2.4) and (2.5) again. For $i \in N_C$ or $j \in N_2$ the processes $x_t^{(i)}$ resp. $x_t^{(j)}$ are Feller branching processes (with immigration). If we condition on these processes, the processes $x_t^{(j)}, j \in N_R$ become independent time-inhomogeneous Brownian motions (with drift), whose distributions are well understood. Thus if the associated process is denoted by $x_t = \left\{x_t^{(j)}\right\}_{j\in N_R \cup N_{C2}} = \left\{x_t^{(j)}\right\}_{j\in V}$, the semigroup $P_t f$ has the explicit representation

$$P_t f(x) = \left(\otimes_{i\in N_{C2}} P^i_{x_i}\right)\left[\int_{\mathbb{R}^{|N_R|}} f\left(\{z_j\}_{j\in N_R}, \left\{x_t^{(i)}\right\}_{i\in N_{C2}}\right) \times \prod_{j\in N_R} p_{\gamma_j^0 2 I_t^{(j)}}\left(z_j - x_j - b_j^0 t\right) dz_j\right], \tag{2.15}$$

where $P^i_{x_i}$ is the law of the Feller branching immigration process $x^{(i)}$ on $\mathcal{C}(\mathbb{R}_+, \mathbb{R}_+)$, started at x_i with generator

$$\mathcal{A}_0^i = b_i^0 \frac{\partial}{\partial x} + \gamma_i^0 x \frac{\partial^2}{\partial x^2},$$
$$I_t^{(j)} = \int_0^t \sum_{i\in C_j} x_s^{(i)} ds,$$

and for $y \in (0,\infty)$

$$p_y(z) := \frac{e^{-\frac{z^2}{2y}}}{(2\pi y)^{1/2}}.$$

Remark 2.2.1. *This also shows that the $\mathcal{A}^0$ martingale problem is well-posed.*

For $(y,z)=\Big(\{y_j\}_{j\in N_R},\{z_i\}_{i\in N_{C2}}\Big)$ and $x^{N_R}\equiv\{x_j\}_{j\in N_R}$, let

$$G(y,z)=G_{t,x^{N_R}}(y,z)=G_{t,x^{N_R}}\Big(\{y_j\}_{j\in N_R},\{z_i\}_{i\in N_{C2}}\Big) \tag{2.16}$$
$$=\int_{\mathbb{R}^{|N_R|}} f\Big(\{u_j\}_{j\in N_R},\{z_i\}_{i\in N_{C2}}\Big)\prod_{j\in N_R} p_{\gamma_j^0 2y_j}\big(u_j-x_j-b_j^0 t\big)\,du_j.$$

Notation 2.2.2. In the following we shall use the notations

$$E^{N_{C2}}=\big(\otimes_{i\in N_{C2}}P^i_{x_i}\big)\,,\ I_t^{N_R}=\Big\{I_t^{(j)}\Big\}_{j\in N_R}\,,\ x_t^{N_{C2}}=\Big\{x_t^{(i)}\Big\}_{i\in N_{C2}}$$

and we shall write E whenever we do not specify w.r.t. which measure we integrate.

Now (2.15) can be rewritten as

$$P_tf(x)=E^{N_{C2}}\Big[G_{t,x^{N_R}}\Big(I_t^{N_R},x_t^{N_{C2}}\Big)\Big]=E^{N_{C2}}\Big[G\Big(I_t^{N_R},x_t^{N_{C2}}\Big)\Big]\,. \tag{2.17}$$

Lemma 2.2.3. *Let $j\in N_R$, then*
(a)

$$E^{N_{C2}}\left[\sum_{i\in C_j}x_t^{(i)}\right]=\sum_{i\in C_j}\left(x_i+b_i^0t\right),$$
$$E^{N_{C2}}\left[\left(\sum_{i\in C_j}x_t^{(i)}\right)^2\right]=\left(\sum_{i\in C_j}x_i\right)^2+\sum_{i\in C_j}\left(2\left(\sum_{k\in C_j}b_k^0+\gamma_i^0\right)x_i\right)t$$
$$+\sum_{i\in C_j}\left(\left(\sum_{k\in C_j}b_k^0+\gamma_i^0\right)b_i^0\right)t^2,$$
$$E^{N_{C2}}\left[\left(\sum_{i\in C_j}\left(x_t^{(i)}-x_i\right)\right)^2\right]=\sum_{i\in C_j}2\gamma_i^0x_it+\sum_{i\in C_j}\left(\left(\sum_{k\in C_j}b_k^0+\gamma_i^0\right)b_i^0\right)t^2$$

and

$$E^{N_{C2}}\left[I_t^{(j)}\right]=E^{N_{C2}}\left[\int_0^t\sum_{i\in C_j}x_s^{(i)}ds\right]=\sum_{i\in C_j}\left(x_it+\frac{b_i^0}{2}t^2\right).$$

(b)

$$E^{N_{C2}}\left[\left(I_t^{(j)}\right)^{-p}\right]\le c(p)t^{-p}\min_{i\in C_j}\left\{(t+x_i)^{-p}\right\}\ \forall p>0.$$

Note. Observe that the requirement $b_i^0>0$ if $i\in(R\cup C)\cap Z$ as in (2.7) is crucial for Lemma 2.2.3(b). As $i\in C_j, j\in N_R$ implies $i\in C\cap Z$, (2.7) guarantees $b_i^0>0$. The bound (b) cannot be applied to $i\in N_2$ in general, as (2.7) only gives $b_i^0\ge 0$ in these cases.

Proof of (a). The first three results follow from Lemma 7(a) in [7] together with the independence of the Feller-diffusions under consideration.

Proof of (b). Proceeding as in the proof of Lemma 7(b) in [7] we obtain

$$\begin{aligned} E^{N_{C2}}\left[\left(I_t^{(j)}\right)^{-p}\right] &\leq c_p e \int_0^\infty E^{N_{C2}}\left[e^{-u^{-1}I_t^{(j)}}\right] u^{-p-1} du \\ &\leq c_p e \min_{i\in C_j}\left\{\int_0^\infty P^i_{x_i}\left[e^{-u^{-1}I_t^{(i)}}\right] u^{-p-1} du\right\} \end{aligned}$$

as $I_t^{(j)} = \sum_{i\in C_j}\int_0^t x_s^{(i)} ds \equiv \sum_{i\in C_j} I_t^{(i)}$, where the Feller-diffusions under consideration are independent. Now we can proceed as in Lemma 7(b) of [7] to obtain the desired result. □

Lemma 2.2.4. *Let $G_{t,x^{N_R}}$ be as in (2.16). Then*

(a) for $j \in N_R$

$$\left|\frac{\partial G_{t,x^{N_R}}}{\partial x_j}\left(\{y_j\}_{j\in N_R}, \{z_i\}_{i\in N_{C2}}\right)\right| = \left|\frac{\partial G_{t,x^{N_R}}}{\partial x_j}(y,z)\right| \leq \| f \|_\infty (\gamma_j^0 y_j)^{-1/2}, \tag{2.18}$$

and more generally for any $k \in \mathbb{N}$, there is a constant c_k such that

$$\left|\frac{\partial^k G_{t,x^{N_R}}}{\partial x_j^k}(y,z)\right| \leq c_k \| f \|_\infty y_j^{-k/2}.$$

(b) For $j \in N_R$

$$\left|\frac{\partial G_{t,x^{N_R}}}{\partial y_j}(y,z)\right| \leq c_1 \| f \|_\infty y_j^{-1}. \tag{2.19}$$

More generally there are constants $c_k, k \in \mathbb{N}$ such that for $l_1, l_2, j_1, j_2 \in N_R$,

$$\left|\frac{\partial^{m_1+m_2+k_1+k_2} G_{t,x^{N_R}}}{\partial x_{l_1}^{m_1}\partial x_{l_2}^{m_2}\partial y_{j_1}^{k_1}\partial y_{j_2}^{k_2}}(y,z)\right| \leq c_{m_1+m_2+k_1+k_2} \| f \|_\infty y_{l_1}^{-m_1/2} y_{l_2}^{-m_2/2} y_{j_1}^{-k_1} y_{j_2}^{-k_2}$$

for all $m_1, m_2, k_1, k_2 \in \mathbb{N}$.

(c) Let $y^{N_R} = \{y_j\}_{j\in N_R}$ and $z^{N_{C2}} = \{z_i\}_{i\in N_{C2}}$, then for all $z^{N_{C2}}$ with $z_i \geq 0$, $i \in N_{C2}$ we have that $\left(x^{N_R}, y^{N_R}\right) \to G_{t,x^{N_R}}\left(y^{N_R}, z^{N_{C2}}\right)$ is $\mathcal{C}^3$ on $\mathbb{R}^{|N_R|} \times (0,\infty)^{|N_R|}$.

Proof. This proceeds as in [7], Lemma 11, using the product form of the density. □

Lemma 2.2.5. *If f is a bounded Borel function on $\mathcal{S}_0$ and $t > 0$, then $P_t f \in \mathcal{C}_b(\mathcal{S}_0)$ with*

$$|P_t f(x) - P_t f(x')| \leq c \| f \|_\infty t^{-1} |x - x'| .$$

Proof. The outline of the proof is as in the proof of [7], Lemma 12. We shall nevertheless show the proof in detail as it illustrates some basic notational issues, which will appear again in later theorems. Note in particular the frequent use of the triangle inequality resulting in additional sums of the form $\sum_{j:j\in\bar{R}_{i_0}}$ in the second part of the proof.

Using (2.17), we have for $x, x' \in \mathbb{R}^{N_R}$,

$$\begin{aligned}
&\left|P_t f\big(x, x^{N_{C2}}\big) - P_t f\big(x', x^{N_{C2}}\big)\right| \qquad (2.20)\\
&= \left|E^{N_{C2}}\left[G_{t,x}\left(I_t^{N_R}, x_t^{N_{C2}}\right) - G_{t,x'}\left(I_t^{N_R}, x_t^{N_{C2}}\right)\right]\right|\\
&\le \|f\|_\infty \sum_{j\in N_R} \frac{|x_j - x_j'|}{\sqrt{\gamma_j^0}} E^{N_{C2}}\left[\left(I_t^{(j)}\right)^{-1/2}\right] \text{ (by (2.18))}\\
&\le c\,\|f\|_\infty \sum_{j\in N_R} \frac{|x_j - x_j'|}{\sqrt{\gamma_j^0}} t^{-1/2} \min_{i\in C_j}\left\{(t+x_i)^{-1/2}\right\} \text{ (by Lemma 2.2.3(b))}\\
&\le c\,\|f\|_\infty\, t^{-1} \sum_{j\in N_R} |x_j - x_j'|.
\end{aligned}$$

Next we shall consider $x, x' = x + he_{i_0} \in \mathbb{R}^{N_{C2}}$ where $i_0 \in N_{C2}$ is arbitrarily fixed. Assume $h > 0$ and let x^h denote an independent copy of $x^{(i_0)}$ starting at h but with $b_{i_0}^0 = 0$. Then $x^{(i_0)} + x^h$ has law $P^{i_0}_{x_{i_0}+h}$ (additive property of Feller branching processes) and so if $I_h(t) = \int_0^t x_s^h ds$,

$$\begin{aligned}
&\left|P_t f\big(x^{N_R}, x'\big) - P_t f\big(x^{N_R}, x\big)\right|\\
&= \Bigg|E^{N_{C2}}\Bigg[G_{t,x^{N_R}}\left(\left\{I_t^{(j)} + 1_{\{i_0\in C_j\}} I_h(t)\right\}_{j\in N_R}, \left\{x_t^i + 1_{\{i=i_0\}} x_t^h\right\}_{i\in N_{C2}}\right)\\
&\qquad -G_{t,x^{N_R}}\left(\left\{I_t^{(j)}\right\}_{j\in N_R}, x_t^{N_{C2}}\right)\Bigg]\Bigg|.
\end{aligned}$$

For what follows it is important to observe that

$$\{j \in N_R : i_0 \in C_j\} = \left\{j : j \in \bar{R}_{i_0}\right\},$$

having made the definition of $\bar{R}_i$ necessary. Next we shall use the triangle inequality to first sum up changes in the jth coordinates (where $j \in N_R$ such that $i_0 \in C_j$) separately in increasing order, followed by the change in the coordinate i_0. If $T_h = \inf\{t \ge 0 : x_t^h = 0\}$ we thus obtain as a bound for the above (recall that e_k denotes the unit vector in the kth direction):

$$\begin{aligned}
&\sum_{j:j\in\bar{R}_{i_0}} c\,\|f\|_\infty\, E^{N_{C2}}\left[I_h(t)\left(I_t^{(j)}\right)^{-1}\right] + 2\,\|f\|_\infty\, E[T_h > t]\\
&= \sum_{j:j\in\bar{R}_{i_0}} c\,\|f\|_\infty\, E^{N_{C2}}[I_h(t)]\, E^{N_{C2}}\left[\left(I_t^{(j)}\right)^{-1}\right] + 2\,\|f\|_\infty\, E[T_h > t]
\end{aligned}$$

by (2.19) and as $\| G \|_\infty \leq \| f \|_\infty$ by the definition of G. Next we shall use that $E[T_h > t] \leq \frac{h}{t\gamma^0_{i_0}}$ (for reference see equation (2.26) in Section 2.2.2). Together with Lemma 2.2.3(a), (b) we may bound the above by

$$\sum_{j: j \in \bar{R}_{i_0}} c \| f \|_\infty \, htt^{-1} \min_{i \in C_j} \{(t + x_i)^{-1}\} + 2 \| f \|_\infty \frac{h}{t\gamma^0_{i_0}} \leq c \| f \|_\infty \, ht^{-1}.$$

The case $x' = x + he_i, i \in N_{C2}$ follows similarly. Note that for $i \in N_2$ only the second term in the above bound is nonzero as the sum is taken over an empty set ($\bar{R}_i = \emptyset$ for $i \in N_2$). Together with (2.20) (recall that the 1-norm and Euclidean norm are equivalent) we obtain the result via triangle inequality. □

Finally, we give elementary calculus inequalities that will be used below.

Lemma 2.2.6. *Let $g : \mathbb{R}^d_+ \to \mathbb{R}$ be $\mathcal{C}^2$. Then for all $\Delta, \Delta' > 0, y \in \mathbb{R}^d_+$ and $I_1, I_2 \subset \{1, \ldots, d\}$,*

$$\begin{aligned}
&\frac{|g(y + \Delta \sum_{i_1 \in I_1} e_{i_1} + \Delta' \sum_{i_2 \in I_2} e_{i_2}) - g(y + \Delta \sum_{i_1 \in I_1} e_{i_1})}{(\Delta\Delta')} \\
&\frac{-g(y + \Delta' \sum_{i_2 \in I_2} e_{i_2}) + g(y)|}{(\Delta\Delta')} \\
&\leq \sup_{\{y' \in \prod_{i \in \{1,\ldots,d\}} [y_i, y_i + \Delta + \Delta']\}} \sum_{i_1 \in I_1} \sum_{i_2 \in I_2} \left| \frac{\partial^2}{\partial y_{i_1} \partial y_{i_2}} g(y') \right|.
\end{aligned}$$

Also let $f : \mathbb{R}^d_+ \to \mathbb{R}$ be $\mathcal{C}^3$. Then for all $\Delta_1, \Delta_2, \Delta_3 > 0, y \in \mathbb{R}^d_+$ and $I_1, I_2, I_3 \subset \{1, \ldots, d\}$,

$$\begin{aligned}
&\frac{|f(y + \Delta_1 \sum_{i_1 \in I_1} e_{i_1} + \Delta_2 \sum_{i_2 \in I_2} e_{i_2} + \Delta_3 \sum_{i_3 \in I_3} e_{i_3})}{(\Delta_1\Delta_2\Delta_3)} \\
&\frac{-f(y + \Delta_1 \sum_{i_1 \in I_1} e_{i_1} + \Delta_3 \sum_{i_3 \in I_3} e_{i_3}) + f(y + \Delta_2 \sum_{i_2 \in I_2} e_{i_2})}{(\Delta_1\Delta_2\Delta_3)} \\
&\frac{-f(y + \Delta_2 \sum_{i_2 \in I_2} e_{i_2} + \Delta_3 \sum_{i_3 \in I_3} e_{i_3}) + f(y + \Delta_3 \sum_{i_3 \in I_3} e_{i_3})}{(\Delta_1\Delta_2\Delta_3)} \\
&\frac{-f(y + \Delta_1 \sum_{i_1 \in I_1} e_{i_1} + \Delta_2 \sum_{i_2 \in I_2} e_{i_2}) + f(y + \Delta_1 \sum_{i_1 \in I_1} e_{i_1}) - f(y)|}{(\Delta_1\Delta_2\Delta_3)} \\
&\leq \sup_{\{y' \in \prod_{i \in \{1,\ldots,d\}} [y_i, y_i + \Delta_1 + \Delta_2 + \Delta_3]\}} \sum_{i_1 \in I_1} \sum_{i_2 \in I_2} \sum_{i_3 \in I_3} \left| \frac{\partial^3}{\partial y_{i_1} \partial y_{i_2} \partial y_{i_3}} f(y') \right|.
\end{aligned}$$

Proof. This is an extension of [7], Lemma 13, using the triangle inequality to split the terms under consideration into sums of differences in only one coordinate at a time. □

2.2.2 Decomposition techniques

In this subsection we cite relevant material from [7], namely Lemma 8, Proposition 9 and Lemma 10. Proofs and references can be found in [7]. Further background and motivation on the processes under consideration may be found in [11], Section II.7.

Let $\{P_x^0 : x \geq 0\}$ denote the laws of the Feller branching process X with no immigration (equivalently, the 0-dimensional squared Bessel process) with generator $\mathcal{L}^0 f(x) = \gamma x f''(x)$. Recall that the Feller branching process X can be constructed as the weak limit of a sequence of rescaled critical Galton-Watson branching processes.

If $\omega \in \mathcal{C}(\mathbb{R}_+, \mathbb{R}_+)$ let $\zeta(\omega) = \inf\{t > 0 : \omega(t) = 0\}$. There is a unique σ-finite measure $\mathbb{N}_0$ on

$$\mathcal{C}_{ex} = \{\omega \in \mathcal{C}(\mathbb{R}_+, \mathbb{R}_+) : \omega(0) = 0, \zeta(\omega) > 0, \omega(t) = 0 \ \forall t \geq \zeta(\omega)\} \tag{2.21}$$

such that for each $h > 0$, if Ξ^h is a Poisson point process on $\mathcal{C}_{ex}$ with intensity $h\mathbb{N}_0$, then

$$X = \int_{\mathcal{C}_{ex}} \nu \Xi^h(d\nu) \text{ has law } P_h^0. \tag{2.22}$$

Citing [11], $\mathbb{N}_0$ can be thought of being the time evolution of a cluster given that it survives for some positive length of time. The representation (2.22) decomposes X according to the ancestors at time 0.

Moreover we also have

$$\mathbb{N}_0[\nu_\delta > 0] = (\gamma\delta)^{-1} \tag{2.23}$$

and for $t > 0$

$$\int_{\mathcal{C}_{ex}} \nu_t d\mathbb{N}_0(\nu) = 1. \tag{2.24}$$

For $t > 0$ let P_t^* denote the probability on $\mathcal{C}_{ex}$ defined by

$$P_t^*[A] = \frac{\mathbb{N}_0[A \cap \{\nu_t > 0\}]}{\mathbb{N}_0[\nu_t > 0]}. \tag{2.25}$$

Lemma 2.2.7. *For all $h > 0$*

$$P_h^0[\zeta > t] = P_h^0[X_t > 0] = 1 - e^{-h/(t\gamma)} \leq \frac{h}{t\gamma}. \tag{2.26}$$

Proposition 2.2.8. *Let $f : \mathcal{C}(\mathbb{R}_+, \mathbb{R}_+) \to \mathbb{R}$ be bounded and continuous. Then for any $\delta > 0$,*

$$\lim_{h \downarrow 0} h^{-1} E_h^0[f(X) 1_{\{X_\delta > 0\}}] = \int_{\mathcal{C}_{ex}} f(\nu) 1_{\{\nu_\delta > 0\}} d\mathbb{N}_0(\nu).$$

The representation (2.22) leads to the following decompositions of the processes $x_t^{(i)}, i \in N_{C2}$ that will be used below. Recall that $x_t^{(i)}$ is the Feller branching immigration process with coefficients $b_i^0 \geq 0, \gamma_i^0 > 0$ starting at x_i and with law $P_{x_i}^i$. In particular, we can make use of the additive property of Feller branching processes.

Lemma 2.2.9. *Let $0 \leq \rho \leq 1$.*

(a) We may assume

$$x^{(i)} = X_0' + X_1,$$

where X_0' is a diffusion with generator $\mathcal{A}_0' f(x) = \gamma_i^0 x f''(x) + b_i^0 f'(x)$ starting at ρx_i, X_1 is a diffusion with generator $\gamma_i^0 x f''(x)$ starting at $(1-\rho)x_i \geq 0$, and X_0', X_1 are independent. In addition, we may assume

$$X_1(t) = \int_{\mathcal{C}_{ex}} \nu_t \Xi(d\nu) = \sum_{j=1}^{N_t} e_j(t), \tag{2.27}$$

where Ξ is a Poisson point process on $\mathcal{C}_{ex}$ with intensity $(1-\rho)x_i\mathbb{N}_0$, $\{e_j, j \in \mathbb{N}\}$ is an iid sequence with common law P_t^*, and N_t is a Poisson random variable (independent of the $\{e_j\}$) with mean $\frac{(1-\rho)x_i}{t\gamma_i^0}$.

(b) We also have

$$\begin{aligned} \int_0^t X_1(s)ds &= \int_{\mathcal{C}_{ex}} \int_0^t \nu_s ds 1_{\{\nu_t \neq 0\}} \Xi(d\nu) + \int_{\mathcal{C}_{ex}} \int_0^t \nu_s ds 1_{\{\nu_t = 0\}} \Xi(d\nu) \\ &\equiv \sum_{j=1}^{N_t} r_j(t) + I_1(t) \end{aligned}$$

and

$$\int_0^t x_s^{(i)} ds = \sum_{j=1}^{N_t} r_j(t) + I_2(t), \tag{2.28}$$

where $r_j(t) = \int_0^t e_j(s)ds, I_2(t) = I_1(t) + \int_0^t X_0'(s)ds$.

(c) Let Ξ^h be a Poisson point process on $\mathcal{C}_{ex}$ with intensity $h_i\mathbb{N}_0$ $(h_i > 0)$, independent of the above processes. Set $\Xi^{x+h} = \Xi + \Xi^h$ and $X_t^h = \int \nu_t \Xi^h(d\nu)$. Then

$$X_t^{x+h} \equiv x_t^{(i)} + X^h(t) = \int_{\mathcal{C}_{ex}} \nu_t \Xi^{x+h}(d\nu) + X_0'(t) \tag{2.29}$$

is a diffusion with generator $\mathcal{A}_0'$ starting at $x_i + h_i$. In addition

$$\int_{\mathcal{C}_{ex}} \nu_t \Xi^{x+h}(d\nu) = \sum_{j=1}^{N_t'} e_j(t), \tag{2.30}$$

where N_t' is a Poisson random variable with mean $((1-\rho)x_i + h_i)(\gamma_i^0 t)^{-1}$, such that $\{e_j\}$ and (N_t, N_t') are independent.

Also

$$\int_0^t X_s^{x+h} ds = \sum_{j=1}^{N_t'} r_j(t) + I_2(t) + I_3^h(t), \tag{2.31}$$

where $I_3^h(t) = \int_{\mathcal{C}_{ex}} \int_0^t \nu_s ds 1_{\{\nu_t = 0\}} \Xi^h(d\nu)$.

2.2.3 Existence and representation of derivatives of the semigroup

Let $\mathcal{A}^0$ and P_t be as in Subsection 2.2.1. The first and second partial derivatives of $P_t f$ w.r.t. $x_k, x_l, k, l \in N_{C2}$ will be represented in terms of the canonical measure $\mathbb{N}_0$.

Recall that by (2.17)

$$P_t f(x) = E^{N_{C2}}\left[G\left(I_t^{N_R}, x_t^{N_{C2}}\right)\right],$$

where $I_t^{N_R} = \left\{I_t^{(j)}\right\}_{j\in N_R}$ with $I_t^{(j)} = \int_0^t \sum_{i\in C_j} x_s^{(i)} ds$.

Notation 2.2.10. If $X \in \mathcal{C}\left(\mathbb{R}_+, \mathbb{R}_+^{N_{C2}}\right)$, $\eta, \eta', \theta, \theta' \in \mathcal{C}_{ex}$ (for the definition of $\mathcal{C}_{ex}$ see (2.21)) and $k, l \in N_{C2}$, let

$$\begin{aligned}
&G^{+k}_{t,x^{N_R}}\left(X; \int_0^t \eta_s ds, \theta_t\right)\\
&\equiv G_{t,x^{N_R}}\left(\left\{\int_0^t \sum_{i\in C_j} X_s^i ds + 1_{\{k\in C_j\}}\int_0^t \eta_s ds\right\}_{j\in N_R}, \left\{X_t^i + 1_{\{i=k\}}\theta_t\right\}_{i\in N_{C2}}\right)
\end{aligned}$$

and

$$\begin{aligned}
&G^{+k,+l}_{t,x^{N_R}}\left(X; \int_0^t \eta_s ds, \theta_t, \int_0^t \eta'_s ds, \theta'_t\right)\\
&\equiv G_{t,x^{N_R}}\left(\left\{\int_0^t \sum_{i\in C_j} X_s^i + 1_{\{k\in C_j\}}\eta_s + 1_{\{l\in C_j\}}\eta'_s ds\right\}_{j\in N_R},\right.\\
&\qquad \left.\left\{X_t^i + 1_{\{i=k\}}\theta_t + 1_{\{i=l\}}\theta'_t\right\}_{i\in N_{C2}}\right).
\end{aligned}$$

Note that if $k \in N_2$ in the above we have $1_{\{k\in C_j\}} = 0$ for $j \in N_R$, i.e.

$$\begin{aligned}
G^{+k}_{t,x^{N_R}}\left(X; \int_0^t \eta_s ds, \theta_t\right) &= G^{+k}_{t,x^{N_R}}\left(X; 0, \theta_t\right),\\
G^{+k,+l}_{t,x^{N_R}}\left(X; \int_0^t \eta_s ds, \theta_t, \int_0^t \eta'_s ds, \theta'_t\right) &= G^{+k,+l}_{t,x^{N_R}}\left(X; 0, \theta_t, \int_0^t \eta'_s ds, \theta'_t\right)
\end{aligned} \tag{2.32}$$

and for $l \in N_2$

$$G^{+k,+l}_{t,x^{N_R}}\left(X; \int_0^t \eta_s ds, \theta_t, \int_0^t \eta'_s ds, \theta'_t\right) = G^{+k,+l}_{t,x^{N_R}}\left(X; \int_0^t \eta_s ds, \theta_t, 0, \theta'_t\right). \tag{2.33}$$

If $X \in \mathcal{C}\left(\mathbb{R}_+, \mathbb{R}_+^{N_{C2}}\right)$, $\nu, \nu' \in \mathcal{C}_{ex}$ and $k, l \in N_{C2}$, let

$$\Delta G^{+k}_{t,x^{N_R}}(X, \nu) \equiv G^{+k}_{t,x^{N_R}}\left(X; \int_0^t \nu_s ds, \nu_t\right) - G^{+k}_{t,x^{N_R}}\left(X; 0, 0\right)$$

and

$$\Delta G^{+k,+l}_{t,x^{N_R}}(X,\nu,\nu') \tag{2.34}$$
$$\equiv G^{+k,+l}_{t,x^{N_R}}\left(X;\int_0^t \nu_s ds,\nu_t,\int_0^t \nu'_s ds,\nu'_t\right) - G^{+k,+l}_{t,x^{N_R}}\left(X;0,0,\int_0^t \nu'_s ds,\nu'_t\right)$$
$$- G^{+k,+l}_{t,x^{N_R}}\left(X;\int_0^t \nu_s ds,\nu_t,0,0\right) + G^{+k,+l}_{t,x^{N_R}}\Big(X;0,0,0,0\Big).$$

Proposition 2.2.11. *If f is a bounded Borel function on $\mathcal{S}_0$ and $t>0$ then $P_t f \in \mathcal{C}^2_b(\mathcal{S}_0)$ and for $k,l \in V = \{1,\ldots,d\}$*

$$\| (P_t f)_{kl} \|_\infty \leq c \frac{\| f \|_\infty}{t^2}.$$

Moreover if f is bounded and continuous on $\mathcal{S}_0$, then for all $k,l \in N_{C2}$

$$(P_t f)_k(x) = E^{N_{C2}}\left[\int \Delta G^{+k}_{t,x^{N_R}}\big(x^{N_{C2}},\nu\big)\, d\mathbb{N}_0(\nu)\right], \tag{2.35}$$

$$(P_t f)_{kl}(x) = E^{N_{C2}}\left[\int\int \Delta G^{+k,+l}_{t,x^{N_R}}\big(x^{N_{C2}},\nu,\nu'\big)\, d\mathbb{N}_0(\nu) d\mathbb{N}_0(\nu')\right]. \tag{2.36}$$

Proof. The outline of this proof is similar to the one for [7], Proposition 14. We shall therefore only mention some changes due to the consideration of more than one catalyst at a time.

With the help of Lemma 2.2.5 and using that $P_t f = P_{t/2}(P_{t/2} f)$ one can easily show that it suffices to consider bounded continuous f. In [7], Proposition 14 one only proves the existence of $(P_t f)_{kl}(x)$, $k,l \in N_{C2}$ and its representation in terms of the canonical measure as in (2.36) based on (2.35). From the methods used it should then be clear how the easier formula (2.35) may have been found.

Hence, let us also assume $(P_t f)_k$ exists and is given by (2.35) for $k \in N_{C2}$. Let $0 < \delta \leq t$. The role of δ will be explained at the end of this proof. In the first case where $\nu'_\delta = \nu_t = 0$, use Lemmas 2.2.6 and 2.2.4(b) to see that for

$k, l \in N_C$

$$
\begin{aligned}
&\left|\Delta G^{+k,+l}_{t,x^{N_R}}\left(x^{N_{C2}}, \nu, \nu'\right)\right| \qquad (2.37)\\
&= \left|G^{+k,+l}_{t,x^{N_R}}\left(x^{N_{C2}}; \int_0^t \nu_s ds, 0, \int_0^\delta \nu'_s ds, 0\right) - G^{+k,+l}_{t,x^{N_R}}\left(x^{N_{C2}}; 0, 0, \int_0^\delta \nu'_s ds, 0\right)\right.\\
&\qquad \left. - G^{+k,+l}_{t,x^{N_R}}\left(x^{N_{C2}}; \int_0^t \nu_s ds, 0, 0, 0\right) + G^{+k,+l}_{t,x^{N_R}}\left(x^{N_{C2}}; 0, 0, 0, 0\right)\right|\\
&= \left|G_{t,x^{N_R}}\left(\left\{\int_0^t \sum_{i \in C_j} x_s^{(i)} ds + 1_{\{k \in C_j\}} \int_0^t \nu_s ds + 1_{\{l \in C_j\}} \int_0^\delta \nu'_s ds\right\}_{j \in N_R}, x_t^{N_{C2}}\right)\right.\\
&\qquad - G_{t,x^{N_R}}\left(\left\{\int_0^t \sum_{i \in C_j} x_s^{(i)} ds + 1_{\{l \in C_j\}} \int_0^\delta \nu'_s ds\right\}_{j \in N_R}, x_t^{N_{C2}}\right)\\
&\qquad - G_{t,x^{N_R}}\left(\left\{\int_0^t \sum_{i \in C_j} x_s^{(i)} ds + 1_{\{k \in C_j\}} \int_0^t \nu_s ds\right\}_{j \in N_R}, x_t^{N_{C2}}\right)\\
&\qquad \left. + G_{t,x^{N_R}}\left(\left\{\int_0^t \sum_{i \in C_j} x_s^{(i)} ds\right\}_{j \in N_R}, x_t^{N_{C2}}\right)\right|\\
&\leq \sum_{j_1 : j_1 \in \bar{R}_k} \sum_{j_2 : j_2 \in \bar{R}_l} c \, \| f \|_\infty \left(I_t^{(j_1)}\right)^{-1} \left(I_t^{(j_2)}\right)^{-1} \int_0^\delta \nu'_s ds \int_0^t \nu_s ds
\end{aligned}
$$

(compare to (49) in [7]).

For k or $l \in N_2$ we obtain via (2.32) and (2.33)

$$\left|\Delta G^{+k,+l}_{t,x^{N_R}}\left(x^{N_{C2}}, \nu, \nu'\right)\right| = 0.$$

This is consistent with (2.37) if we consider the sum over an empty set to be zero (recall that $\bar{R}_k = R_k \cap N_R$ and thus $\bar{R}_k = \emptyset$ if $k \in N_2$). Hence (2.37) is a bound for all $k, l \in N_{C2}$.

The other cases are proven as in [7] (for the last case use the trivial bound $\left|\Delta G^{+k,+l}_{t,x^{N_R}}(x^{N_{C2}}, \nu, \nu')\right| \leq 4 \, \| f \|_\infty$) with the same modifications as just observed.

Combining all the cases we conclude that

$$\left|\Delta G^{+k,+l}_{t,x^{N_R}}(x^{N_{C2}},\nu,\nu')\right|$$
$$\leq \Bigg\{ 1_{\{\nu'_\delta=\nu_t=0\}} \left(\sum_{j_1:j_1\in\bar{R}_k} \sum_{j_2:j_2\in\bar{R}_l} \left(I_t^{(j_1)}\right)^{-1} \left(I_t^{(j_2)}\right)^{-1} \int_0^\delta \nu'_s ds \int_0^t \nu_s ds \right)$$
$$+ 1_{\{\nu'_\delta=0,\nu_t>0\}} \left(\sum_{j:j\in\bar{R}_l} \left(I_t^{(j)}\right)^{-1} \int_0^\delta \nu'_s ds \right)$$
$$+ 1_{\{\nu'_\delta>0,\nu_t=0\}} \left(\sum_{j:j\in\bar{R}_k} \left(I_t^{(j)}\right)^{-1} \int_0^t \nu_s ds \right) + 1_{\{\nu'_\delta>0,\nu_t>0\}} \Bigg\} c \parallel f \parallel_\infty$$
$$\leq \Bigg\{ 1_{\{\nu'_\delta=\nu_t=0\}} \left(\int_0^t x_s^{(k)} ds\right)^{-1} \left(\int_0^t x_s^{(l)} ds\right)^{-1} \int_0^\delta \nu'_s ds \int_0^t \nu_s ds$$
$$+ 1_{\{\nu'_\delta=0,\nu_t>0\}} \left(\int_0^t x_s^{(l)} ds\right)^{-1} \int_0^\delta \nu'_s ds$$
$$+ 1_{\{\nu'_\delta>0,\nu_t=0\}} \left(\int_0^t x_s^{(k)} ds\right)^{-1} \int_0^t \nu_s ds + 1_{\{\nu'_\delta>0,\nu_t>0\}} \Bigg\} c \parallel f \parallel_\infty$$
$$\equiv \bar{g}_{t,\delta}\big(x^{N_{C2}},\nu,\nu'\big)$$

The remainder of the proof works similar to the proof in [7]. Some minor changes are necessary in the proof of continuity from below in x_2 (now to be replaced by $x^{N_{C2}}$) following (59) in [7], by considering every coordinate on its own. Also, new mixed partial derivatives appear, which can be treated similarly to the ones already appearing in the proof of Proposition 14 in [7]. Other necessary technical changes will reappear in later proofs where they will be worked out in detail. They are thus omitted at this point. □

Remark 2.2.12. *The necessity for introducing δ only becomes clear in the context of a complete proof. For instance, the derivation of (2.36) starts by defining $X^h_\cdot$, independent of $x^{(l)}$ and satisfying*

$$X_t^h = h + \int_0^t \sqrt{2\gamma_l^0 X_s^h} dB'_s, \ (h>0)$$

(i.e. X^h has law P_h^0) so that $x^{(l)} + X^h$ has law $P^l_{x_l+h}$. Therefore (2.35) together with definition (2.34) implies

$$\frac{1}{h}\left[(P_t f)_k(x+he_l) - (P_t f)_k(x)\right]$$
$$= \frac{1}{h} \int\int\int \Delta G^{+k,+l}_{t,x^{N_R}}\left(x^{N_{C2}},\nu,X^h\right)\left(1_{\{X^h_\delta=0\}} + 1_{\{X^h_\delta>0\}}\right) d\mathbb{N}_0(\nu) dP^{N_{C2}} dP_h^0.$$

Now the first term can be made arbitrarily small for t fixed and $\delta \downarrow 0^+$. The second term can be further rewritten with the help of Proposition 2.2.8 and will finally yield the representation (2.36) by first taking $h \downarrow 0^+$ and then $\delta \downarrow 0^+$.

2.2.4 L^∞ bounds of certain differentiation operators applied to $P_t f$ and equivalence of norms

We continue to work with the semigroup P_t on the state space $\mathcal{S}_0$ corresponding to the generator $\mathcal{A}^0$. Recall the definitions of the semigroup norm $|f|_\alpha$ from (2.12) and of the associated Banach space of functions S^α from (2.13) in what follows.

Proposition 2.2.13. *If f is a bounded Borel function on $\mathcal{S}_0$ then for $j \in N_R$*

$$\left|\frac{\partial}{\partial x_j} P_t f(x)\right| \leq \frac{c \parallel f \parallel_\infty}{\sqrt{t} \max\limits_{i \in C_j}\{\sqrt{t + x_i}\}}, \tag{2.38}$$

and

$$\left|\max_{i \in C_j}\{x_i\} \frac{\partial^2}{\partial x_j^2} P_t f(x)\right| \leq \frac{c \parallel f \parallel_\infty}{t}. \tag{2.39}$$

If $f \in \mathcal{S}^\alpha$, then

$$\left|\frac{\partial}{\partial x_j} P_t f(x)\right| \leq \frac{c|f|_\alpha t^{\frac{\alpha}{2} - \frac{1}{2}}}{\max\limits_{i \in C_j}\{\sqrt{t + x_i}\}} \leq c|f|_\alpha t^{\frac{\alpha}{2} - 1}, \tag{2.40}$$

and

$$\left|\max_{i \in C_j}\{x_i\} \frac{\partial^2}{\partial x_j^2} P_t f(x)\right| \leq c|f|_\alpha t^{\frac{\alpha}{2} - 1}. \tag{2.41}$$

Proof. The proof proceeds as in [7], Proposition 16 except for minor changes.

The estimate in (2.38) can be obtained by mimicking the calculation in (2.20). (2.39) follows from a double application of (2.38), where we use that P_t and $\frac{\partial}{\partial x_j}$ commute.

If $f \in \mathcal{S}^\alpha$, we proceed as in [2] and write

$$\left|\frac{\partial}{\partial x_j} P_{2t} f(x) - \frac{\partial}{\partial x_j} P_t f(x)\right| = \left|\frac{\partial}{\partial x_j} P_t (P_t f - f)(x)\right|.$$

Applying the estimate (2.38) to $g = P_t f - f$ and using the definition of $|f|_\alpha$ we get

$$\left|\frac{\partial}{\partial x_j} P_{2t} f(x) - \frac{\partial}{\partial x_j} P_t f(x)\right| \leq \frac{c \parallel g \parallel_\infty}{\sqrt{t} \max\limits_{i \in C_j}\{\sqrt{t + x_i}\}} \leq \frac{c|f|_\alpha t^{\alpha/2}}{\sqrt{t} \max\limits_{i \in C_j}\{\sqrt{t + x_i}\}}.$$

This together with

$$(2.38) \Rightarrow \lim_{t \to \infty} \left|\frac{\partial}{\partial x_j} P_t f(x)\right| = 0$$

implies that

$$\begin{aligned}\left|\frac{\partial}{\partial x_j}P_tf(x)\right| &\le \sum_{k=0}^{\infty}\left|\frac{\partial}{\partial x_j}\left(P_{2^kt}f - P_{2^{(k+1)}t}f\right)(x)\right| \\ &\le |f|_\alpha \sum_{k=0}^{\infty}\left(2^kt\right)^{\frac{\alpha}{2}-\frac{1}{2}}\frac{c}{\max_{i\in C_j}\{\sqrt{2^kt+x_i}\}} \\ &\le |f|_\alpha t^{\frac{\alpha}{2}-\frac{1}{2}}\frac{c}{\max_{i\in C_j}\{\sqrt{t+x_i}\}}.\end{aligned}$$

This then immediately yields (2.40). Use (2.39) to derive (2.41) in the same way as (2.38) was used to prove (2.40). □

Notation 2.2.14. If $w > 0$, set $p_j(w) = \frac{w^j}{j!}e^{-w}$. For $\{r_j(t)\}$ and $\{e_j(t)\}$ as in Lemma 2.2.9, let $R_k = R_k(t) = \sum_{j=1}^k r_j(t)$ and $S_k = S_k(t) = \sum_{j=1}^k e_j(t)$.

Notation 2.2.15. If $X \in \mathcal{C}\left(\mathbb{R}_+, \mathbb{R}_+^{N_{C2}}\right)$, $Y, Y', Z, Z' \in \mathcal{C}(\mathbb{R}_+, \mathbb{R}_+)$, η, η', $\theta, \theta' \in \mathcal{C}_{ex}$ and $m, n, k, l \in N_{C2}$, where $m \neq n$ let

$$\begin{aligned}&G^{m,n,+k,+l}_{t,x^{N_R}}\left(X, Y_t, Z_t, Y'_t, Z'_t; \int_0^t \eta_s ds, \theta_t, \int_0^t \eta'_s ds, \theta'_t\right) \\ &\equiv G_{t,x^{N_R}}\Bigg(\Bigg\{\int_0^t \sum_{i\in C_j\setminus\{m,n\}} X^i_s ds + 1_{\{m\in C_j\}}Y_t + 1_{\{n\in C_j\}}Y'_t \\ &\qquad\qquad + \int_0^t 1_{\{k\in C_j\}}\eta_s + 1_{\{l\in C_j\}}\eta'_s ds\Bigg\}_{j\in N_R}, \\ &\qquad \Bigg\{1_{\{i\notin\{m,n\}\}}X^i_t + 1_{\{i=m\}}Z_t + 1_{\{i=n\}}Z'_t \\ &\qquad\qquad + 1_{\{i=k\}}\theta_t + 1_{\{i=l\}}\theta'_t\Bigg\}_{i\in N_{C2}}\Bigg).\end{aligned}$$

The notation indicates that the one-dimensional coordinate processes $\int_0^t X^m_s ds, X^m_t$ resp. $\int_0^t X^n_s ds, X^n_t$ will be replaced by the processes Y_t, Z_t resp. Y'_t, Z'_t (note that for $m \in N_2$ this only implies a change from X^m_t into Z_t). Additionally, we add $\int_0^t \nu_s ds, \theta_t, \int_0^t \nu'_s ds$ and θ'_t as before. The terms

$$G^{m,+k,+l}_{t,x^{N_R}}, G^{m,+k}_{t,x^{N_R}}, G^{m,n,+l}_{t,x^{N_R}}, G^{m,n}_{t,x^{N_R}}, G^{m}_{t,x^{N_R}}, \Delta G^{m,+k,+l}_{t,x^{N_R}} \text{ etc.} \tag{2.42}$$

will then be defined in a similar way, where for instance $G^m_{t,x^{N_R}}$ only refers to replacing the processes $\int_0^t X^m_s ds, X^m_t$ via Y_t, Z_t but doesn't involve adding processes.

Proposition 2.2.16. *If f is a bounded Borel function on $\mathcal{S}_0$, then for $i \in N_{C2}$*

$$\left|\frac{\partial}{\partial x_i} P_t f(x)\right| \leq \frac{c \, \| f \|_\infty}{\sqrt{t}\sqrt{t+x_i}}, \tag{2.43}$$

and

$$\left|x_i \frac{\partial^2}{\partial x_i^2} P_t f(x)\right| \leq \frac{c x_i \, \| f \|_\infty}{t(t+x_i)} \leq \frac{c \, \| f \|_\infty}{t}. \tag{2.44}$$

If $f \in \mathcal{S}^\alpha$, then

$$\left|\frac{\partial}{\partial x_i} P_t f(x)\right| \leq \frac{c|f|_\alpha t^{\frac{\alpha}{2}-\frac{1}{2}}}{\sqrt{t+x_i}} \leq c|f|_\alpha t^{\frac{\alpha}{2}-1},$$

and

$$\left|x_i \frac{\partial^2}{\partial x_i^2} P_t f(x)\right| \leq c|f|_\alpha t^{\frac{\alpha}{2}-1}.$$

Proof. The outline of the proof is the same as for [7], Proposition 17. Part of the proof will be presented here with its notational modifications since some care is needed when working in a multi-dimensional setting and the formulas become more involved.

As in the proof of Proposition 2.2.11 we assume w.l.o.g. that f is bounded and continuous. In what follows we shall illustrate the proof of (2.44) as (2.43) is easier. Consider second derivatives in k. The representation of $(P_t f)_{kk}$ in Proposition 2.2.11 and symmetry allow us to write for $k \in N_{C2}$ (i.e. $l = k$)

$$\begin{aligned}
(P_t f)_{kk}(x) = {} & E^{N_{C2}}\left[\int\int \Delta G^{+k,+k}_{t,x^{N_R}}\left(x^{N_{C2}}, \nu, \nu'\right) 1_{\{\nu_t=0,\nu'_t=0\}} d\mathbb{N}_0(\nu) d\mathbb{N}_0(\nu')\right] \\
& + 2E^{N_{C2}}\left[\int\int \Delta G^{+k,+k}_{t,x^{N_R}}\left(x^{N_{C2}}, \nu, \nu'\right) 1_{\{\nu_t=0,\nu'_t>0\}} d\mathbb{N}_0(\nu) d\mathbb{N}_0(\nu')\right] \\
& + E^{N_{C2}}\left[\int\int \Delta G^{+k,+k}_{t,x^{N_R}}\left(x^{N_{C2}}, \nu, \nu'\right) 1_{\{\nu_t>0,\nu'_t>0\}} d\mathbb{N}_0(\nu) d\mathbb{N}_0(\nu')\right] \\
& \equiv E_1 + 2E_2 + E_3.
\end{aligned}$$

The idea for bounding $|E_1|, |E_2|$ and $|E_3|$ is similar to the one in [7]. In what follows we shall illustrate the necessary changes to bound $|E_3|$.

Notation 2.2.17. We have $\mathbb{N}_0[\cdot \cap \{\nu_t > 0\}] = (\gamma t)^{-1} P_t^*[\cdot]$ on $\{\nu_t > 0\}$, where we used (2.25) and (2.23). Whenever we change integration w.r.t. $\mathbb{N}_0$ to integration w.r.t. P_t^* we shall denote this by $\stackrel{(*)}{=}$.

The decomposition of Lemma 2.2.9 (cf. (2.27) and (2.28)) with $\rho = 0$ gives

$$
\begin{aligned}
|E_3| \overset{(*)}{=} \frac{c}{t^2}\Big| E\Big[\int\int\Big\{ & G^{k,+k,+k}_{t,x^{N_R}}\Big(x^{N_{C2}}, R_{N_t} + I_2(t), S_{N_t} + X_0'(t); \\
& \qquad\qquad \int_0^t \nu_s ds, \nu_t, \int_0^t \nu_s' ds, \nu_t'\Big) \\
& - G^{k,+k,+k}_{t,x^{N_R}}\Big(x^{N_{C2}}, R_{N_t} + I_2(t), S_{N_t} + X_0'(t); 0, 0, \int_0^t \nu_s' ds, \nu_t'\Big) \\
& - G^{k,+k,+k}_{t,x^{N_R}}\Big(x^{N_{C2}}, R_{N_t} + I_2(t), S_{N_t} + X_0'(t); \int_0^t \nu_s ds, \nu_t, 0, 0\Big) \\
& + G^{k,+k,+k}_{t,x^{N_R}}\Big(x^{N_{C2}}, R_{N_t} + I_2(t), S_{N_t} + X_0'(t); 0, 0, 0, 0\Big)\Big\} \\
& \times dP_t^*(\nu) dP_t^*(\nu')\Big]\Big|,
\end{aligned} \tag{2.45}
$$

where for instance

$$
\begin{aligned}
& G^{k,+k,+k}_{t,x^{N_R}}\Big(x^{N_{C2}}, R_{N_t} + I_2(t), S_{N_t} + X_0'(t); \int_0^t \nu_s ds, \nu_t, \int_0^t \nu_s' ds, \nu_t'\Big) \\
& = G_{t,x^{N_R}}\Bigg(\Bigg\{\int_0^t \sum_{i \in C_j \setminus \{k\}} X_s^i ds + 1_{\{k \in C_j\}}\left(R_{N_t} + I_2(t)\right) \\
& \qquad\qquad + \int_0^t 1_{\{k \in C_j\}}\left(\nu_s + \nu_s'\right) ds\Bigg\}_{j \in N_R}, \\
& \qquad \Big\{1_{\{i \neq k\}} X_t^i + 1_{\{i = k\}}\left(S_{N_t} + X_0'(t)\right) \\
& \qquad\qquad + 1_{\{i = k\}}\left(\nu_t + \nu_t'\right)\Big\}_{i \in N_{C2}}\Bigg)
\end{aligned}
$$

by Notation 2.2.15 and the comment following it.

Recall that $R_k = R_k(t) = \sum_{j=1}^k r_j(t)$ and $S_k = S_k(t) = \sum_{j=1}^k e_j(t)$ with $\{r_j(t)\}$ and $\{e_j(t)\}$ as in Lemma 2.2.9. In particular, $\{e_j, j \in \mathbb{N}\}$ is iid with common law P_t^* and $r_j(t) = \int_0^t e_j(s) ds$.

We obtain (recall the definition of $G^k_{t,x^{N_R}}$ from (2.42))

$$
\begin{aligned}
|E_3| = \frac{c}{t^2}\Big| E\Big[& G^k_{t,x^{N_R}}\Big(x^{N_{C2}}, R_{N_t+2} + I_2(t), S_{N_t+2} + X_0'(t)\Big) \\
& - 2G^k_{t,x^{N_R}}\left(x^{N_{C2}}, R_{N_t+1} + I_2(t), S_{N_t+1} + X_0'(t)\right) \\
& + G^k_{t,x^{N_R}}\Big(x^{N_{C2}}, R_{N_t} + I_2(t), S_{N_t} + X_0'(t)\Big)\Big]\Big|.
\end{aligned}
$$

Observe that in case $k \in N_2$ the above notation $G^k_{t,x^{N_R}}(x^{N_{C2}}, R_{N_t} + I_2(t), S_{N_t} + X_0'(t))$ only indicates that $x_t^{(k)}$ gets changed into $S_{N_t} + X_0'(t)$; for $k \in N_2$

the indicated change of $\int_0^t x_s^{(k)} ds$ into $R_{N_t} + I_2(t)$ has no impact on the term under consideration.

Let $w = x_k/(\gamma_k^0 t)$. The independence of N_t from $(\{\int_0^t x_s^{(i)} ds, i \in C_j \backslash \{k\}, j \in N_R\}, x_t^{(N_{C2})\backslash\{k\}}, I_2(t), X_0'(t), \{e_l\}, \{r_l\})$ yields

$$
\begin{aligned}
|E_3| = \frac{c}{t^2}\Big|\sum_{n=0}^{\infty} p_n(w) E\Big[& G^k_{t,x^{N_R}}\big(x^{N_{C2}}, R_{n+2} + I_2(t), S_{n+2} + X_0'(t)\big) \\
& - 2G^k_{t,x^{N_R}}\big(x^{N_{C2}}, R_{n+1} + I_2(t), S_{n+1} + X_0'(t)\big) \\
& + G^k_{t,x^{N_R}}\big(x^{N_{C2}}, R_n + I_2(t), S_n + X_0'(t)\big)\Big]\Big|.
\end{aligned}
$$

Sum by parts twice and use $|G| \leq \| f \|_\infty$ to bound the above by

$$
\begin{aligned}
& c \| f \|_\infty \frac{1}{x_k t}\left(w(3p_0(w) + p_1(w)) + \sum_{n=2}^{\infty} w \, |p_{n-2}(w) - 2p_{n-1}(w) + p_n(w)| \right) \\
& \leq c \| f \|_\infty \frac{1}{x_k t}\left(wp_0(w) + wp_1(w) + \sum_{n=2}^{\infty} p_n(w) \frac{|(w-n)^2 - n|}{w} \right) \\
& \leq c \| f \|_\infty \frac{1}{x_k t}\left(2p_1(w) + \sum_{n=0}^{\infty} p_n(w) \frac{(w-n)^2 + n}{w} \right) \\
& \leq c \| f \|_\infty \frac{1}{x_k t}.
\end{aligned}
$$

We obtain another bound on $|E_3|$ if we use the trivial bound $|G| \leq \| f \|_\infty$ in (2.45). This yields $|E_3| \leq c \| f \|_\infty t^{-2}$ and so

$$
|E_3| \leq \frac{c \| f \|_\infty}{t(t + x_k)}.
$$

Combine the bounds on $|E_1|, |E_2|$ and $|E_3|$ to obtain (2.44).

The bounds for $f \in \mathcal{S}^\alpha$ are obtained from the above just as in the proof of Proposition 2.2.13. □

Recall Convention 2.1.11, as stated in (2.14), for the definition of M^0 in what follows.

Notation 2.2.18. Set $J_t^{(j)} = \gamma_j^0 2 I_t^{(j)}, j \in N_R$.

Lemma 2.2.19. *For each $M \geq 1$, $\alpha \in (0,1)$ and $d \in \mathbb{N}$ there is a $c = c(M, \alpha, d) > 0$ such that if $M^0 \leq M$, then*

$$
|fg|_\alpha \leq c|f|_{\mathcal{C}_w^\alpha} \| g \|_\infty + \| f \|_\infty |g|_\alpha \tag{2.46}
$$

and

$$
\| fg \|_\alpha \leq c \left(\| f \|_{\mathcal{C}_w^\alpha} \| g \|_\infty + \| f \|_\infty |g|_\alpha \right). \tag{2.47}
$$

Proof. Compared to the proof of [7], Lemma 18, the derivation of a bound for the second error term E_2 below becomes more involved. Again the triangle-inequality has to be used to express multi-dimensional coordinate changes via one-dimensional ones.

Let $\left(x^{N_R}, x^{N_{C2}}\right) \in \mathbb{R}^{|N_R|} \times \mathbb{R}_+^{|N_{C2}|}$ and define $\tilde{f}(y) = f(y) - f(x)$. Then (2.15) gives

$$\begin{aligned} & |P_t(fg)(x) - (fg)(x)| \qquad (2.48)\\ & \leq |P_t(\tilde{f}g)(x)| + |f(x)||P_t g(x) - g(x)| \\ & \leq \|g\|_\infty \, E^{N_{C2}}\left[\int_{\mathbb{R}^{|N_R|}} \left|\tilde{f}\left(z^{N_R}, x_t^{N_{C2}}\right)\right| \prod_{j\in N_R} p_{J_t^{(j)}}\left(z_j - x_j - b_j^0 t\right) dz_j\right] \\ & \quad + \|f\|_\infty \, |g|_\alpha t^{\alpha/2}. \end{aligned}$$

The above expectation can be bounded by three terms as follows:

$$\begin{aligned} & E^{N_{C2}}\left[\int \left|\tilde{f}\left(z^{N_R}, x_t^{N_{C2}}\right)\right| \prod_{j\in N_R} p_{J_t^{(j)}}\left(z_j - x_j - b_j^0 t\right) dz_j\right] \qquad (2.49)\\ & \leq E^{N_{C2}}\Bigg[\int \Big\{\left|\tilde{f}\left(z^{N_R}, x_t^{N_{C2}}\right) - \tilde{f}\left(z^{N_R}, x^{N_{C2}}\right)\right| \\ & \qquad + \left|f\left(z^{N_R}, x^{N_{C2}}\right) - f\left(x^{N_R} + b_{N_R}^0 t, x^{N_{C2}}\right)\right| \\ & \qquad + \left|f\left(x^{N_R} + b_{N_R}^0 t, x^{N_{C2}}\right) - f\left(x^{N_R}, x^{N_{C2}}\right)\right|\Big\} \\ & \qquad \times \prod_{j\in N_R} p_{J_t^{(j)}}\left(z_j - x_j - b_j^0 t\right) dz_j\Bigg] \\ & \equiv E_1 + E_2 + E_3. \end{aligned}$$

For all three terms we shall use the triangle inequality to sum up changes in different coordinates separately.

The definition of $|f|_{\alpha,i}$ gives

$$\begin{aligned} E_1 & \leq \sum_{i\in N_{C2}} |f|_{\alpha,i} E^{N_{C2}}\left[\left(\left|x_t^{(i)} - x_i\right|^\alpha x_i^{-\alpha/2}\right) \wedge \left|x_t^{(i)} - x_i\right|^{\alpha/2}\right] \\ & \leq \sum_{i\in N_{C2}} |f|_{\alpha,i}\left(\left(E^{N_{C2}}\left[\left|x_t^{(i)} - x_i\right|^2\right]^{\alpha/2} x_i^{-\alpha/2}\right) \wedge E^{N_{C2}}\left[\left|x_t^{(i)} - x_i\right|^2\right]^{\alpha/4}\right). \end{aligned}$$

We now proceed as in the derivation of a bound on E_1 in the proof of Lemma 18 in [7], using Lemma 2.2.3(a) (alternatively compare with estimation of E_2 below). We finally obtain

$$E_1 \leq c \sum_{i\in N_{C2}} |f|_{\alpha,i} t^{\alpha/2} 2^{\alpha/2} \leq c|f|_{\mathcal{C}_w^\alpha} t^{\alpha/2} 2^{\alpha/2}.$$

Similarly we have

$$E_2 \le \sum_{k\in N_R} \min_{i:k\in\bar{R}_i} \Big\{ |f|_{\alpha,i} E^{N_{C2}} \Big[\int \Big(\Big(\big| z_k - (x_k + b_k^0 t) \big|^{\alpha} x_i^{-\alpha/2} \Big) \wedge$$

$$\big| z_k - (x_k + b_k^0 t)\big|^{\alpha/2} \Big) \prod_{j\in N_R} p_{J_t^{(j)}} (z_j - x_j - b_j^0 t)\, dz_j \Big] \Big\}$$

$$\le c \sum_{k\in N_R} \min_{i:k\in\bar{R}_i} \Big\{ |f|_{\alpha,i} E^{N_{C2}} \Big[\Big(\Big| J_t^{(k)} \Big|^{\alpha/2} x_i^{-\alpha/2} \Big) \wedge \Big| J_t^{(k)} \Big|^{\alpha/4} \Big] \Big\}$$

$$\le c \sum_{k\in N_R} \min_{i:k\in\bar{R}_i} \Big\{ |f|_{\alpha,i} \Big(\Big(E^{N_{C2}} \Big[\Big| J_t^{(k)} \Big| \Big]^{\alpha/2} x_i^{-\alpha/2} \Big) \wedge E^{N_{C2}} \Big[\Big| J_t^{(k)} \Big| \Big]^{\alpha/4} \Big) \Big\}$$

as $\int |z|^{\beta} p_J(z) dz \le c J^{\beta/2}$ for $\beta \in (0,1)$. Next use Lemma 2.2.3(a) which shows that $E^{N_{C2}}\Big[J_t^{(k)}\Big] = \gamma_k^0 2 E^{N_{C2}}\Big[I_t^{(k)}\Big] \le \sum_{l\in C_k} cM^2(t^2 + x_l t)$. Put this in the above bound on E_2 to see that E_2 can be bounded by

$$c \sum_{k\in N_R} \min_{i:k\in\bar{R}_i} \Bigg\{ |f|_{\alpha,i} \Bigg(\Bigg(\Bigg(\sum_{l\in C_k} (t^2 + x_l t) \Bigg)^{\alpha/2} x_i^{-\alpha/2} \Bigg) \wedge \Bigg(\sum_{l\in C_k} (t^2 + x_l t) \Bigg)^{\alpha/4} \Bigg) \Bigg\}$$

$$\overset{k\in N_R}{\le} c|f|_{\mathcal{C}_w^\alpha} \sum_{k\in N_R} \Bigg(\Bigg(\sum_{l\in C_k} \frac{t^2 + x_l t}{\max\limits_{i:k\in\bar{R}_i} x_i} \Bigg)^{\alpha/2} \wedge \Bigg(\sum_{l\in C_k} \Big(t^2 + t \max_{i:k\in\bar{R}_i} x_i \Big) \Bigg)^{\alpha/4} \Bigg)$$

$$\overset{k\in N_R}{\le} c|f|_{\mathcal{C}_w^\alpha} t^{\alpha/2} \sum_{k\in N_R} \Bigg(\Bigg(\frac{t}{\max\limits_{i:k\in\bar{R}_i} x_i} + 1 \Bigg)^{\alpha/2} \wedge \Bigg(1 + \frac{\max\limits_{i:k\in\bar{R}_i} x_i}{t} \Bigg)^{\alpha/4} \Bigg)$$

$$\le\ c|f|_{\mathcal{C}_w^\alpha} t^{\alpha/2} 2^{\alpha/2}.$$

For the third term E_3 we finally have

$$E_3 \le \sum_{k\in N_R} \min_{i:k\in\bar{R}_i} \Big\{ |f|_{\alpha,i} \Big(\Big(\big| b_k^0 t \big|^{\alpha} x_i^{-\alpha/2} \Big) \wedge \Big(\big| b_k^0 t \big|^{\alpha/2} \Big) \Big) \Big\}$$

$$\le c|f|_{\mathcal{C}_w^\alpha} \sum_{k\in N_R} \big| b_k^0 t \big|^{\alpha/2}$$

$$\le c|f|_{\mathcal{C}_w^\alpha} t^{\alpha/2}.$$

Put the above bounds on E_1, E_2 and E_3 into (2.49) and then in (2.48) to conclude that

$$|P_t(fg)(x) - (fg)(x)| \le \big(\| g \|_\infty\, c|f|_{\mathcal{C}_w^\alpha} + \| f \|_\infty\, |g|_\alpha \big) t^{\alpha/2}$$

and so by definition of the semigroup norm

$$|fg|_\alpha \le c|f|_{\mathcal{C}_w^\alpha}\, \| g \|_\infty + \| f \|_\infty\, |g|_\alpha.$$

This gives (2.46) and (2.47) is then immediate. □

Theorem 2.2.20. *There exist* $0 < c_1 \le c_2$ *such that*

$$c_1 |f|_{\mathcal{C}^\alpha_w} \le |f|_\alpha \le c_2 |f|_{\mathcal{C}^\alpha_w}. \tag{2.50}$$

This implies that $\mathcal{C}^\alpha_w = \mathcal{S}^\alpha$ *and so* $\mathcal{S}^\alpha$ *contains* $\mathcal{C}^1$ *functions with compact support in* $\mathcal{S}_0$.

Proof. The idea of the proof was taken from the proof of Theorem 19 in [7]. The second inequality in (2.50) follows immediately by setting $g = 1$ in Lemma 2.2.19. For the first inequality let $x, h \in \mathcal{S}_0$, $t > 0$ and use Propositions 2.2.13 and 2.2.16 to see that

$$\begin{aligned} &|f(x+h) - f(x)| \qquad (2.51)\\ &\le |P_t f(x+h) - f(x+h)| + |P_t f(x) - f(x)| + |P_t f(x+h) - P_t f(x)| \\ &\le 2|f|_\alpha t^{\alpha/2} + |P_t f(x+h) - P_t f(x)| \\ &\le 2|f|_\alpha t^{\alpha/2} + c|f|_\alpha t^{\frac{\alpha}{2}-\frac{1}{2}} \left(\sum_{j \in N_R} \frac{|h_j|}{\max_{l \in C_j}\{\sqrt{t+x_l}\}} + \sum_{i \in N_{C2}} \frac{h_i}{\sqrt{t+x_i}} \right), \end{aligned}$$

where we used the triangle inequality together with $h_l \ge 0, l \in C_j \subset N_{C2}$ for all $j \in N_R$.

By setting $t = |h|$ and bounding $\left(\max_{l \in C_j}\{\sqrt{t+x_l}\}\right)^{-1}$ and $\left(\sqrt{t+x_i}\right)^{-1}$ by $\left(\sqrt{t}\right)^{-1}$ we obtain as a first bound on (2.51)

$$c|f|_\alpha |h|^{\alpha/2}. \tag{2.52}$$

Next only consider $h \in \mathcal{S}_0$ such that there exists $i \in N_{C2}$ and $j \in \{i\} \cup \bar{R}_i$ such that $h_j \neq 0$ and $h_k = 0$ if $k \notin \{i\} \cup \bar{R}_i$. (2.51) becomes

$$\begin{aligned} &|f(x+h) - f(x)| \\ &\le 2|f|_\alpha t^{\alpha/2} + c|f|_\alpha t^{\frac{\alpha}{2}-\frac{1}{2}} \left(\sum_{j : j \in \bar{R}_i} \frac{|h_j|}{\max_{l \in C_j}\{\sqrt{t+x_l}\}} + \frac{h_i}{\sqrt{t+x_i}} \right) \\ &\le 2|f|_\alpha t^{\alpha/2} + c|f|_\alpha t^{\frac{\alpha}{2}-\frac{1}{2}} \frac{1}{\sqrt{t+x_i}} |h|. \end{aligned}$$

In case $x_i > 0$ set $t = \frac{|h|^2}{x_i}$ and bound $\left(\sqrt{t+x_i}\right)^{-1}$ by $\left(\sqrt{x_i}\right)^{-1}$ to get as a second upper bound

$$c|f|_\alpha x_i^{-\alpha/2} |h|^\alpha. \tag{2.53}$$

The first inequality in (2.50) is now immediate from (2.52) and (2.53) and the proof is complete. □

Note. Special care was needed when choosing $h \in \mathcal{S}_0$ in the last part of the proof as it only works for those h which are to be considered in the definition of $|\cdot|_{\mathcal{C}^\alpha_w}$. Note that this was the main reason to define the weighted Hölder norms for $\bar{R}_i$ instead of R_i.

Remark 2.2.21. *The equivalence of the two norms will prove to be crucial later in Section 2.3, where we show the uniqueness of solutions to the martingale problem MP($\mathcal{A}$,ν) as stated in Theorem 2.1.6. All the estimates of Section 2.2 are obtained in terms of the semigroup norm. In Section 2.3 we shall further need estimates on the norm of products of certain functions. At this point we shall have to rely on the result of Lemma 2.2.19 for weighted Hölder norms. The equivalence of norms now yields a similar result in terms of the semigroup norm.*

2.2.5 Weighted Hölder bounds of certain differentiation operators applied to $P_t f$

The $x_j, j \in N_R$ derivatives are much easier.

Notation 2.2.22. We shall need the following slight extension of our notation for $E^{N_{C2}}$:

$$E^{N_{C2}} = E^{N_{C2}}_{x^{N_{C2}}} = \left(\otimes_{i\in N_{C2}} P^i_{x_i}\right).$$

Notation 2.2.23. To ease notation let

$$T_k^{-\frac{1}{2}}\left(t, x^{N_{C2}}\right) \equiv \begin{cases} \min\limits_{l\in C_k}\left\{(t+x_l)^{-1/2}\right\}, & k \in N_R, \\ (t+x_k)^{-1/2}, & k \in N_{C2}. \end{cases}$$

Proposition 2.2.24. *If f is a bounded Borel function on $\mathcal{S}_0$, then for all $x, h \in \mathcal{S}_0$, $j \in N_R$, $i \in C_j$ and arbitrary $k \in V$,*

$$\left|\frac{\partial}{\partial x_j} P_t f(x + h_k e_k) - \frac{\partial}{\partial x_j} P_t f(x)\right| \leq \frac{c\,\|f\|_\infty}{t^{3/2}} |h_k| T_k^{-\frac{1}{2}}\left(t, x^{N_{C2}}\right) \tag{2.54}$$

and

$$\left|(x + h_k e_k)_i \frac{\partial^2 P_t f}{\partial x_j^2}(x + h_k e_k) - x_i \frac{\partial^2 P_t f}{\partial x_j^2}(x)\right| \leq \frac{c\,\|f\|_\infty}{t^{3/2}} |h_k| T_k^{-\frac{1}{2}}\left(t, x^{N_{C2}}\right). \tag{2.55}$$

If $f \in \mathcal{S}^\alpha$, then

$$\left|\frac{\partial}{\partial x_j} P_t f(x + h_k e_k) - \frac{\partial}{\partial x_j} P_t f(x)\right| \leq c|f|_\alpha t^{\frac{\alpha}{2}-\frac{3}{2}} |h_k| T_k^{-\frac{1}{2}}\left(t, x^{N_{C2}}\right) \tag{2.56}$$

and

$$\left|(x + h_k e_k)_i \frac{\partial^2 P_t f}{\partial x_j^2}(x + h_k e_k) - x_i \frac{\partial^2 P_t f}{\partial x_j^2}(x)\right| \leq c|f|_\alpha t^{\frac{\alpha}{2}-\frac{3}{2}} |h_k| T_k^{-\frac{1}{2}}\left(t, x^{N_{C2}}\right). \tag{2.57}$$

Proof. The focus will be on proving (2.55) as (2.54) is simpler. Again, it suffices to consider f bounded and continuous. For increments in x_k, $k \in N_R$ the statement follows as in the proof of [7], Proposition 22.

Consider increments in x_k, $k \in N_{C2}$. We start with observing that for $h_k \geq 0$

$$\begin{aligned}
&(x_i + \delta_{ki}h_i)\frac{\partial^2 P_t f}{\partial x_j^2}(x + h_k e_k) - x_i \frac{\partial^2 P_t f}{\partial x_j^2}(x) \\
&= \delta_{ki}h_i E^{N_{C2}}_{x^{N_{C2}}+h_k e_k}\left[\frac{\partial^2}{\partial x_j^2} G_{t,x^{N_R}}\left(I_t^{N_R}, x_t^{N_{C2}}\right)\right] \\
&\quad + x_i\left(E^{N_{C2}}_{x^{N_{C2}}+h_k e_k} - E^{N_{C2}}_{x^{N_{C2}}}\right)\left[\frac{\partial^2}{\partial x_j^2} G_{t,x^{N_R}}\left(I_t^{N_R}, x_t^{N_{C2}}\right)\right] \\
&\equiv E_1 + E_2,
\end{aligned}$$

by arguing as in the proof of [7], Proposition 22. The bound on E_1 is derived as in that proof, using Lemmas 2.2.4(a) and 2.2.3(b).

For E_2 we use the decompositions (2.29), (2.30), (2.31) and notation from Lemma 2.2.9 with $\rho = \frac{1}{2}$. Recall the notation $G^k_{t,x^{N_R}}$ from (2.42) and the definition of R_k and S_k as in Notation 2.2.14. Then

$$\begin{aligned}
|E_2| &= x_i \left| E\left[\frac{\partial^2}{\partial x_j^2} G^k_{t,x^{N_R}}\left(x^{N_{C2}}, R_{N'_t} + I_2(t) + I_3^h(t), S_{N'_t} + X'_0(t)\right)\right.\right. \\
&\qquad \left.\left. - \frac{\partial^2}{\partial x_j^2} G^k_{t,x^{N_R}}\left(x^{N_{C2}}, R_{N_t} + I_2(t), S_{N_t} + X'_0(t)\right)\right]\right| \\
&\leq x_i \left| E\left[\frac{\partial^2}{\partial x_j^2} G^k_{t,x^{N_R}}\left(x^{N_{C2}}, R_{N'_t} + I_2(t) + I_3^h(t), S_{N'_t} + X'_0(t)\right)\right.\right. \\
&\qquad \left.\left. - \frac{\partial^2}{\partial x_j^2} G^k_{t,x^{N_R}}\left(x^{N_{C2}}, R_{N'_t} + I_2(t), S_{N'_t} + X'_0(t)\right)\right]\right| \\
&\quad + x_i \left| E\left[\frac{\partial^2}{\partial x_j^2} G^k_{t,x^{N_R}}\left(x^{N_{C2}}, R_{N'_t} + I_2(t), S_{N'_t} + X'_0(t)\right)\right.\right. \\
&\qquad \left.\left. - \frac{\partial^2}{\partial x_j^2} G^k_{t,x^{N_R}}\left(x^{N_{C2}}, R_{N_t} + I_2(t), S_{N_t} + X'_0(t)\right)\right]\right| \\
&\equiv E_{2a} + E_{2b}.
\end{aligned}$$

E_{2a} can be bounded as in [7], using Lemmas 2.2.4(b) and 2.2.3(b), and the independence of $x^{N_{C2}}$ and $I_3^h(t)$. Next turn to E_{2b}. Recall that $S_n = S_n(t) = \sum_{l=1}^n e_l(t)$, $R_n = R_n(t) = \sum_{l=1}^n r_l(t)$ and $p_k(w) = e^{-w}w^k/k!$. In the first term of E_{2b} we may condition on N'_t as it is independent from the other random variables and in the second term we do the same for N_t. Thus, if $w' = w + \frac{h_k}{\gamma_k^0 t}$

and $w = \frac{x_k}{2\gamma_k^0 t}$, then by Lemma 2.2.4(a) and Lemma 2.2.3(b),

$$\begin{aligned}
&E_{2b}\\
&= x_i \left| \sum_{n=0}^{\infty} (p_n(w') - p_n(w)) \, E\left[\frac{\partial^2}{\partial x_j^2} G^k_{t,x^{N_R}} \left(x^{N_{C2}}, R_n + I_2(t), S_n + X_0'(t)\right)\right]\right| \\
&\le c x_i \sum_{n=0}^{\infty} \left| \int_w^{w'} p_n'(u) du \right| \|f\|_\infty \\
&\quad \times \begin{cases} E^{N_{C2}}\left[\left(I_t^{(j)}\right)^{-1}\right], & k \notin C_j, \\ \min\limits_{i \in C_j \setminus \{k\}} \left\{ E^{N_{C2}}\left[\left(\int_0^t x_s^{(i)} ds\right)^{-1}\right]\right\} \wedge E\left[\left(\int_0^t X_0'(s) ds\right)^{-1}\right], & k \in C_j \end{cases} \\
&\le c \|f\|_\infty \, x_i \sum_{n=0}^{\infty} \left| \int_w^{w'} p_n'(u) du \right| t^{-1} \min_{l \in C_j} \left\{ (t + x_l)^{-1} \right\},
\end{aligned}$$

where we used that X_0' starts at $\frac{x_k}{2}$ and thus by Lemma 2.2.3(b)

$$E\left[\left(\int_0^t X_0'(s) ds\right)^{-1}\right] \le ct^{-1} \left(t + \frac{x_k}{2}\right)^{-1} \le ct^{-1}(t + x_k)^{-1}.$$

We therefore obtain with $i \in C_j$

$$\begin{aligned}
E_{2b} &\le c \|f\|_\infty \, x_i \int_w^{w'} \sum_{n=0}^{\infty} p_n(u) \frac{|n-u|}{u} du t^{-1} (t + x_i)^{-1} \\
&\le c \|f\|_\infty \left(\left(\int_w^{w'} \frac{1}{\sqrt{u}} du \right) \wedge \left(\int_w^{w'} 2 du \right) \right) t^{-1},
\end{aligned}$$

where we used $\sum_{n=0}^{\infty} p_n(u) \frac{|n-u|}{u} = \frac{1}{u} E|N-u| \le \frac{1}{u}\sqrt{E|N-u|^2} = \frac{1}{\sqrt{u}}$ and $\sum_{n=0}^{\infty} p_n(u) \frac{|n-u|}{u} \le \sum_{n=0}^{\infty} p_n(u) \left(\frac{n}{u} + 1\right) = \frac{E|N|}{u} + 1 = 2$ with N being Poisson distributed with parameter u. Hence

$$E_{2b} \le c \|f\|_\infty \, (w' - w) \left(\frac{1}{\sqrt{w}} \wedge 2 \right) t^{-1} = c \|f\|_\infty \frac{h_k}{t} \left(\frac{\sqrt{t}}{\sqrt{x_k}} \wedge 2 \right) t^{-1}.$$

As $\left(\frac{1}{\sqrt{x_k}} \wedge \frac{2}{\sqrt{t}}\right) \le c \frac{1}{\sqrt{t + x_k}}$ we finally get

$$E_{2b} \le c \|f\|_\infty \, t^{-3/2} h_k (t + x_k)^{-1/2}.$$

The bounds (2.56) and (2.57) can be derived from the first two by an argument similar to the one used in the proof of Proposition 2.2.13 (alternatively refer to the end of the proof of Proposition 22 in [7]). □

In what follows recall Notation 2.2.23.

Proposition 2.2.25. *If f is a bounded Borel function on $\mathcal{S}_0$, then for all $x, h \in \mathcal{S}_0$, $i \in N_{C2}$ and arbitrary $k \in V$,*

$$\left|\frac{\partial}{\partial x_i}P_t f(x+h_k e_k) - \frac{\partial}{\partial x_i}P_t f(x)\right| \leq \frac{c\,\|f\|_\infty}{t^{3/2}}|h_k| T_k^{-\frac{1}{2}}\left(t, x^{N_{C2}}\right) \tag{2.58}$$

and

$$\left|(x+h_k e_k)_i \frac{\partial^2 P_t f}{\partial x_i^2}(x+h_k e_k) - x_i \frac{\partial^2 P_t f}{\partial x_i^2}(x)\right| \leq \frac{c\,\|f\|_\infty}{t^{3/2}}|h_k| T_k^{-\frac{1}{2}}\left(t, x^{N_{C2}}\right). \tag{2.59}$$

If $f \in \mathcal{S}^\alpha$, then

$$\left|\frac{\partial}{\partial x_i}P_t f(x+h_k e_k) - \frac{\partial}{\partial x_i}P_t f(x)\right| \leq c|f|_\alpha t^{\frac{\alpha}{2}-\frac{3}{2}}|h_k| T_k^{-\frac{1}{2}}\left(t, x^{N_{C2}}\right)$$

and

$$\left|(x+h_k e_k)_i \frac{\partial^2 P_t f}{\partial x_i^2}(x+h_k e_k) - x_i \frac{\partial^2 P_t f}{\partial x_i^2}(x)\right| \leq c|f|_\alpha t^{\frac{\alpha}{2}-\frac{3}{2}}|h_k| T_k^{-\frac{1}{2}}\left(t, x^{N_{C2}}\right).$$

Proof. Proposition 2.2.25 is an extension of Proposition 23 in [7]. The last two inequalities follow from the first two by an argument similar to the one used in the proof of Proposition 2.2.13 (alternatively refer to the end of the proof of Proposition 22 in [7]). As the proof of (2.58) is similar to, but much easier than, that of (2.59), we only prove the latter. As usual we may assume f is bounded and continuous.

Recall the notation $\Delta G^{+i,+i}_{t,x^{N_R}}(X,\nu,\nu')$ from (2.34). Proposition 2.2.11 gives

$$(P_t f)_{ii}(x) = \sum_{n=1}^{4} E^{N_{C2}}\left[\Delta_n G_{t,x^{N_R}}\left(x^{N_{C2}}\right)\right], \tag{2.60}$$

where

$$\Delta_1 G_{t,x^{N_R}}(X) \equiv \int\int \Delta G^{+i,+i}_{t,x^{N_R}}(X,\nu,\nu') 1_{\{\nu_t=\nu'_t=0\}} d\mathbb{N}_0(\nu) d\mathbb{N}_0(\nu'),$$

$$\Delta_2 G_{t,x^{N_R}}(X) \equiv \int\int \Delta G^{+i,+i}_{t,x^{N_R}}(X,\nu,\nu') 1_{\{\nu_t>0,\nu'_t=0\}} d\mathbb{N}_0(\nu) d\mathbb{N}_0(\nu'),$$

$$\Delta_3 G_{t,x^{N_R}}(X) \equiv \int\int \Delta G^{+i,+i}_{t,x^{N_R}}(X,\nu,\nu') 1_{\{\nu_t=0,\nu'_t>0\}} d\mathbb{N}_0(\nu) d\mathbb{N}_0(\nu')$$

and

$$\begin{aligned}\Delta_4 G_{t,x^{N_R}}(X) &\equiv \int\int \Delta G^{+i,+i}_{t,x^{N_R}}(X,\nu,\nu') 1_{\{\nu_t>0,\nu'_t>0\}} d\mathbb{N}_0(\nu) d\mathbb{N}_0(\nu') \\ &\overset{(*)}{=} \frac{c}{t^2}\int\int \Delta G^{+i,+i}_{t,x^{N_R}}(X,\nu,\nu') 1_{\{\nu_t>0,\nu'_t>0\}} dP_t^*(\nu) dP_t^*(\nu').\end{aligned}$$

Let us consider first the increments in $x_k, k \in N_{C2}$. Increments in $x_k, k \in N_R$ will follow at the end of this section in Lemma 2.2.30. Let $h_k \geq 0$ and use (2.60) to obtain

$$\begin{aligned} & |(x+h_k e_k)_i (P_t f)_{ii}(x+h_k e_k) - x_i (P_t f)_{ii}(x)| \\ & \leq \sum_{n=1}^{4} \left| x_i \left(E^{N_{C2}}_{x^{N_{C2}}+h_k e_k} - E^{N_{C2}}_{x^{N_{C2}}} \right) \left[\Delta_n G_{t,x^{N_R}} \left(x^{N_{C2}} \right) \right] \right| \\ & \quad + h_k \left| (P_t f)_{kk}(x+h_k e_k) \right| . \end{aligned} \tag{2.61}$$

The last term on the right hand side can be bounded via (2.44) as follows:

$$h_k \left| (P_t f)_{kk}(x+h_k e_k) \right| \leq h_k \frac{c \, \| f \|_\infty}{t(t+x_k)} \leq c \, \| f \|_\infty \, h_k t^{-3/2} (t+x_k)^{-1/2},$$

where we used $h_k \geq 0$.

In the following Lemmas 2.2.26, 2.2.27 and 2.2.29 we again use the decompositions from Lemma 2.2.9 with $\rho = \frac{1}{2}$ to bound the first four terms in (2.61).

Lemma 2.2.26. *For $k \in N_{C2}$ (and $i \in N_{C2}$) we have*

$$\left| x_i \left(E^{N_{C2}}_{x^{N_{C2}}+h_k e_k} - E^{N_{C2}}_{x^{N_{C2}}} \right) \left[\Delta_1 G_{t,x^{N_R}} \left(x^{N_{C2}} \right) \right] \right| \leq \frac{c \, \| f \|_\infty}{t^{3/2} (t+x_k)^{1/2}} h_k.$$

Proof. This Lemma corresponds to Lemma 24 in [7]. In [7] one considered $\Delta G^{+i,+i}(\cdot)$ as a second order difference, thus obtaining terms involving $(t+x_i)^{-2}$. In our setting this method will not work for $i \neq k$ as we do in fact need terms of the form $(t+x_i)^{-1}(t+x_k)^{-1}$. Instead, we shall bound the left hand side by reasoning as for the E_2-term in Proposition 22 of [7] (part of the proof can be found in this paper in the proof of Proposition 2.2.24), but with $\frac{\partial^2}{\partial x_j^2} G(\cdot), j \in N_R$ replaced by $\Delta G^{+i,+i}(\cdot), i \in N_{C2}$. □

Lemma 2.2.27. *For $k \in N_{C2}$ (and $i \in N_{C2}$) and $n = 2, 3$ we have*

$$\left| x_i \left(E^{N_{C2}}_{x^{N_{C2}}+h_k e_k} - E^{N_{C2}}_{x^{N_{C2}}} \right) \left[\Delta_n G_{t,x^{N_R}} \left(x^{N_{C2}} \right) \right] \right| \leq \frac{c \, \| f \|_\infty}{t^{3/2} (t+x_k)^{1/2}} h_k. \tag{2.62}$$

Proof. By symmetry we only need to consider $n = 2$. As before let $w = \frac{x_k}{2\gamma_k^0 t}$, $w' = w + \frac{h_k}{\gamma_k^0 t}$, $S_n = \sum_{l=1}^n e_l(t)$ and $R_n = \sum_{l=1}^n r_l(t)$. Let Q_h be the law of $I_3^h(t)$ as defined after (2.31). As this random variable is independent of the others appearing below we may condition on it and use (2.29), (2.30) and (2.31) to conclude

$$
\begin{aligned}
&x_i E^{N_{C2}}_{x^{N_{C2}}+h_k e_k}\left[\Delta_2 G_{t,x^{N_R}}\left(x^{N_{C2}}\right)\right]\\
&= x_i E\Bigg[\int\int\int\Bigg\{G^{k,+i,+i}_{t,x^{N_R}}\Bigg(x^{N_{C2}}, I_2(t)+z+R_{N_t'}, X_0'(t)+S_{N_t'};\\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad \int_0^t \nu_s ds, \nu_t, \int_0^t \nu_s' ds, 0\Bigg)\\
&\quad - G^{k,+i,+i}_{t,x^{N_R}}\Bigg(x^{N_{C2}}, I_2(t)+z+R_{N_t'}, X_0'(t)+S_{N_t'}; 0, 0, \int_0^t \nu_s' ds, 0\Bigg)\\
&\quad - G^{k,+i,+i}_{t,x^{N_R}}\Bigg(x^{N_{C2}}, I_2(t)+z+R_{N_t'}, X_0'(t)+S_{N_t'}; \int_0^t \nu_s ds, \nu_t, 0, 0\Bigg)\\
&\quad + G^{k,+i,+i}_{t,x^{N_R}}\Big(x^{N_{C2}}, I_2(t)+z+R_{N_t'}, X_0'(t)+S_{N_t'}; 0, 0, 0, 0\Big)\Bigg\}\\
&\qquad\quad \times 1_{\{\nu_t>0\}} 1_{\{\nu_t'=0\}} d\mathbb{N}_0(\nu) d\mathbb{N}_0(\nu') dQ_h(z)\Bigg].
\end{aligned}
$$

When working under $E^{N_{C2}}_{x^{N_{C2}}}$ there is no $I_3^h(t)$ term. Hence we obtain the same formula with z replaced by 0 and N_t' replaced by N_t. The difference of these terms can be bounded by a difference dealing with the change from z to 0 and the change from N_t' to N_t separately. For the second term we recall that $p_n(u) = e^{-u}u^n/n!$ and observe that N_t' is independent of the other random variables. Hence we may condition on its value to see that the l.h.s. of (2.62) is at most

$$
\begin{aligned}
&x_i \Bigg| E\Bigg[\int\int\int\Bigg\{\Delta G^{k,+i,+i}_{t,x^{N_R}}\Big(x^{N_{C2}}, I_2(t)+z+R_{N_t'}, X_0'(t)+S_{N_t'};\\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad \int_0^t \nu_s ds, \nu_t, \int_0^t \nu_s' ds, 0\Big)\\
&\quad - \Delta G^{k,+i,+i}_{t,x^{N_R}}\Bigg(x^{N_{C2}}, I_2(t)+R_{N_t'}, X_0'(t)+S_{N_t'}; \int_0^t \nu_s ds, \nu_t, \int_0^t \nu_s' ds, 0\Bigg)\Bigg\}\\
&\qquad \times 1_{\{\nu_t>0\}} 1_{\{\nu_t'=0\}} d\mathbb{N}_0(\nu) d\mathbb{N}_0(\nu') dQ_h(z)\Bigg]\Bigg|\\
&\quad + x_i \Bigg|\sum_{n=0}^{\infty} (p_n(w') - p_n(w)) E\Bigg[\int\int \Delta G^{k,+i,+i}_{t,x^{N_R}}\Big(x^{N_{C2}},\\
&\qquad\qquad\qquad I_2(t)+R_n, X_0'(t)+S_n; \int_0^t \nu_s ds, \nu_t, \int_0^t \nu_s' ds, 0\Big)\\
&\qquad\qquad\qquad \times 1_{\{\nu_t>0\}} 1_{\{\nu_t'=0\}} d\mathbb{N}_0(\nu) d\mathbb{N}_0(\nu')\Bigg]\Bigg|\\
&\equiv E_a + E_b.
\end{aligned}
$$

The first term can be rewritten as the sum of two second order differences (one in z, one in $\int_0^t \nu_s' ds$). Together with Lemma 2.2.6, Lemma 2.2.4(b) and

Lemma 2.2.3(b) we therefore obtain (terms including empty sums are again understood as being zero)

$$
\begin{aligned}
E_a \le\ & 2x_i c \,\|f\|_\infty \sum_{j_1 : j_1 \in \bar{R}_i} \sum_{j_2 : j_2 \in \bar{R}_k} \\
& E\left[\begin{cases} \left(I_t^{(j_1)}\right)^{-1}, & k \notin C_{j_1}, \\ \min_{m \in C_{j_1} \setminus \{k\}} \left\{ \left(\int_0^t x_s^{(m)} ds\right)^{-1} \right\} \wedge \left(\int_0^t X_0'(s) ds\right)^{-1}, & k \in C_{j_1} \end{cases} \right. \\
& \quad \left. \times \left(\int_0^t X_0'(s) ds \right)^{-1} \right] \int \int_0^t \nu_s' ds d\mathbb{N}_0(\nu') \mathbb{N}_0[\nu_t > 0] \int z dQ_h(z) \\
\le\ & x_i c \,\|f\|_\infty\, t^{-2} (t + x_i)^{-1} (t + x_k)^{-1} t t^{-1} h_k t \\
\le\ & c \,\|f\|_\infty\, h_k t^{-3/2} (t + x_k)^{-1/2}.
\end{aligned}
$$

Turning to E_b observe that we have the sum of two first order differences (both in $\int_0^t \nu_s' ds$). Together with the triangle inequality, Lemma 2.2.4(b) and Lemma 2.2.3(b) we therefore obtain

$$
\begin{aligned}
E_b \le\ & c x_i \sum_{n=0}^{\infty} \left| \int_w^{w'} p_n'(u) du \right| \|f\|_\infty \sum_{j_1 : j_1 \in \bar{R}_i} \\
& E\left[\begin{cases} \left(I_t^{(j_1)}\right)^{-1}, & k \notin C_{j_1}, \\ \min_{m \in C_{j_1} \setminus \{k\}} \left\{ \left(\int_0^t x_s^{(m)} ds\right)^{-1} \right\} \wedge \left(\int_0^t X_0'(s) ds\right)^{-1}, & k \in C_{j_1} \end{cases} \right] \\
& \times \int \int_0^t \nu_s' ds d\mathbb{N}_0(\nu') \mathbb{N}_0[\nu_t > 0] \\
\le\ & c x_i \sum_{n=0}^{\infty} \left| \int_w^{w'} p_n'(u) du \right| \|f\|_\infty\, t^{-1} (t + x_i)^{-1} t t^{-1}.
\end{aligned}
$$

Now proceed again as in the estimation of E_{2b} in the proof of Proposition 2.2.24 to get

$$
\begin{aligned}
E_b &\le c x_i t^{-1/2} h_k (t + x_k)^{-1/2} \,\|f\|_\infty\, t^{-1} (t + x_i)^{-1} t t^{-1} \\
&\le c \,\|f\|_\infty\, h_k t^{-3/2} (t + x_k)^{-1/2}.
\end{aligned}
$$

The above bounds on E_a and E_b give the required result. □

Notation 2.2.28. Let

$$
G_{t, x^{N_R}}^{m, n \neq m}(X, Y_t, Z_t, Y_t', Z_t') \equiv \begin{cases} G_{t, x^{N_R}}^{m,n}(X, Y_t, Z_t, Y_t', Z_t') & \text{if } n \neq m \\ G_{t, x^{N_R}}^{m}(X, Y_t, Z_t) & \text{if } n = m. \end{cases}
$$

Expressions such as $G_{t, x^{N_R}}^{m, n \neq m, +k, +l}\left(X, Y_t, Z_t, Y_t', Z_t'; \int_0^t \eta_s ds, \theta_t, \int_0^t \eta_s' ds, \theta_t'\right)$ will be defined similarly.

Lemma 2.2.29. *For $k \in N_{C2}$ (and $i \in N_{C2}$) we have*

$$\left|x_i \left(E^{N_{C2}}_{x^{N_{C2}}+h_k e_k} - E^{N_{C2}}_{x^{N_{C2}}}\right)\left[\Delta_4 G_{t,x^{N_R}}\left(x^{N_{C2}}\right)\right]\right| \le \frac{c \parallel f \parallel_\infty}{t^{3/2}(t+x_k)^{1/2}} h_k.$$

Proof. Let

$$E \equiv x_i \left|\left(E^{N_{C2}}_{x^{N_{C2}}+h_k e_k} - E^{N_{C2}}_{x^{N_{C2}}}\right)\left[\Delta_4 G_{t,x^{N_R}}\left(x^{N_{C2}}\right)\right]\right|. \tag{2.63}$$

We use the same setting and notation as in Lemma 2.2.27. Proceeding as in the estimation of the l.h.s. in (2.62), thereby not only decomposing $x^{(k)}$ but also $x^{(i)}$ (the respective parts of the decomposition of $x^{(k)}$ and $x^{(i)}$ are designated via upper indices k resp. i and are independent for $k \neq i$), we have

$$\begin{aligned}
&x_i E^{N_{C2}}_{x^{N_{C2}}+h_k e_k}\left[\Delta_4 G_{t,x^{N_R}}\left(x^{N_{C2}}\right)\right] \\
&= x_i E\Bigg[\int\int\int \Delta G^{k,i\neq k,+i,+i}_{t,x^{N_R}}\Bigg(x^{N_{C2}}, I_2^{(k)}(t) + z + R^{(k)}_{N_t^{'(k)}}, X_0^{'(k)}(t) \\
&\qquad + S^{(k)}_{N_t^{'(k)}}, I_2^{(i)}(t) + R^{(i)}_{N_t^{(i)}}, X_0^{'(i)}(t) + S^{(i)}_{N_t^{(i)}}; \int_0^t \nu_s ds, \nu_t, \int_0^t \nu'_s ds, \nu'_t\Bigg) \\
&\qquad \times 1_{\{\nu_t>0\}} 1_{\{\nu'_t>0\}} d\mathbb{N}_0(\nu) d\mathbb{N}_0(\nu') dQ_h(z)\Bigg].
\end{aligned}$$

Now let for $k = i$

$$\hat{G}_n(z) \equiv E\Big[G^k_{t,x^{N_R}}\Big(x^{N_{C2}}, I_2^{(k)}(t) + z + R_n^{(k)}, X_0^{'(k)}(t) + S_n^{(k)}\Big)\Big],$$

respectively for $k \neq i$,

$$\begin{aligned}
\hat{G}_n\Big(z, N_t^{'(k)}\Big) \equiv E\Big[G^{k,i}_{t,x^{N_R}}\Big(&x^{N_{C2}}, I_2^{(k)}(t) + z + R^{(k)}_{N_t^{'(k)}}, X_0^{'(k)}(t) + S^{(k)}_{N_t^{'(k)}}, \\
&I_2^{(i)}(t) + R_n^{(i)}, X_0^{'(i)}(t) + S_n^{(i)}\Big)\Big].
\end{aligned}$$

Note that the expectation in the definition of $\hat{G}_n\Big(z, N_t^{'(k)}\Big)$ excludes the random variable $N_t^{'(k)}$. Use $w^{'(k)} = \frac{x_k}{2\gamma_k^0 t} + \frac{h_k}{\gamma_k^0 t}$ (i.e. $\rho = 1/2$) to obtain for $k = i$

$$\begin{aligned}
&x_k E^{N_{C2}}_{x^{N_{C2}}+h_k e_k}\left[\Delta_4 G_{t,x^{N_R}}\left(x^{N_{C2}}\right)\right] \\
&\overset{(*)}{=} c\frac{x_k}{t^2}\sum_{n=0}^{\infty} p_n\Big(w^{'(k)}\Big)\int\Big(\hat{G}_{n+2} - 2\hat{G}_{n+1} + \hat{G}_n\Big)(z) dQ_h(z),
\end{aligned} \tag{2.64}$$

and use $w^{(i)} = \frac{x_i}{2\gamma_i^0 t}$ to obtain for $k \neq i$

$$\begin{aligned}
&x_i E^{N_{C2}}_{x^{N_{C2}}+h_k e_k}\left[\Delta_4 G_{t,x^{N_R}}\left(x^{N_{C2}}\right)\right] \\
&\overset{(*)}{=} c\frac{x_i}{t^2}\sum_{n=0}^{\infty} p_n\Big(w^{(i)}\Big) E\Big[\int\Big(\hat{G}_{n+2} - 2\hat{G}_{n+1} + \hat{G}_n\Big)\Big(z, N_t^{'(k)}\Big) dQ_h(z)\Big].
\end{aligned} \tag{2.65}$$

A similar argument holds for $x_i E^{N_{C2}}_{x^{N_{C2}}}\left[\Delta_4 G_{t,x^{N_R}}\left(x^{N_{C2}}\right)\right]$. Indeed, if $k = i$ replace z by 0 and replace $w^{'(k)}$ by $w^{(k)} = \frac{x_k}{2\gamma_k^0 t}$ in (2.64). If $k \neq i$ replace z by 0 and replace $N_t^{'(k)}$ by $N_t^{(k)}$ in (2.65).

Let us first investigate the case $k = i$. Define

$$\hat{H}_n(z) = \hat{G}_n(z) - \hat{G}_n(0)$$

to get for E as in (2.63),

$$\begin{aligned} E &\leq c\frac{x_k}{t^2}\left|\sum_{n=0}^{\infty} p_n\left(w^{'(k)}\right)\int\left(\hat{H}_{n+2} - 2\hat{H}_{n+1} + \hat{H}_n\right)(z)dQ_h(z)\right| \\ &\quad + c\frac{x_k}{t^2}\left|\sum_{n=0}^{\infty}\left(p_n\left(w^{'(k)}\right) - p_n\left(w^{(k)}\right)\right)\left(\hat{G}_{n+2} - 2\hat{G}_{n+1} + \hat{G}_n\right)(0)\right| \\ &\equiv E_1 + E_2. \end{aligned}$$

We can bound E_1 by

$$c\frac{x_k}{t^2}\sum_{n=0}^{\infty}\left|(p_{n-2} - 2p_{n-1} + p_n)\left(w^{'(k)}\right)\right|\sup_{n\geq 0}\left|\int \hat{H}_n(z)dQ_h(z)\right|,$$

where $p_n(w) \equiv 0$ if $n < 0$. By using $q_n(w) = wp_n(w)$ and $\sum_{n=0}^{\infty}|(q_{n-2} - 2q_{n-1} + q_n)(w)| \leq 2$ (see [7], (109)) we obtain

$$E_1 \leq c\frac{x_k}{t^2}\frac{1}{w^{'(k)}}\sup_{n\geq 0}\left|\int \hat{H}_n(z)dQ_h(z)\right|.$$

Next observe that $\hat{H}_n(z)$ is zero for $k \in N_2$ (recall that for $k \in N_2$ the indicated change from $\int_0^t x_s^{(k)}ds$ into $I_2^{(k)}(t) + z + R_n^{(k)}$ resp. $I_2^{(k)}(t) + R_n^{(k)}$ has no impact on the terms under consideration) and is a first order difference for $k \in N_C$ for which we obtain as usual

$$\begin{aligned} \left|\int \hat{H}_n(z)dQ_h(z)\right| &\leq c\,\|f\|_\infty\, t^{-1}(t + x_k)^{-1}\int z dQ_h(z) \\ &\leq c\,\|f\|_\infty\, t^{-1}(t + x_k)^{-1}h_k t \\ &\leq c\,\|f\|_\infty\, h_k t^{-1/2}(t + x_k)^{-1/2}. \end{aligned}$$

Together with $w^{(k)} = \frac{x_k}{2\gamma_k^0 t}$ and $w^{'(k)} = \frac{x_k}{2\gamma_k^0 t} + \frac{h_k}{\gamma_k^0 t}$ this gives

$$E_1 \leq c\,\|f\|_\infty\, h_k t^{-3/2}(t + x_k)^{-1/2}.$$

For E_2 we obtain with $\| G \|_\infty \leq \| f \|_\infty$ and Fubini's theorem

$$E_2 \leq c \, \| f \|_\infty \, \frac{x_k}{t^2} \sum_{n=0}^{\infty} \Big| (p_{n-2} - 2p_{n-1} + p_n) \Big(w^{'(k)} \Big) - \, (p_{n-2} - 2p_{n-1} + p_n) \Big(w^{(k)} \Big) \Big| \leq c \, \| f \|_\infty \, \frac{x_k}{t^2} \int_{w^{(k)}}^{w^{'(k)}} \sum_{n=0}^{\infty} \big| \big(p'_{n-2} - 2p'_{n-1} + p'_n\big)(u) \big| \, du. \tag{2.66}$$

As $p_n(u) = e^{-u} \frac{u^n}{n!}$ we have $p'_n(u) = -p_n(u) + p_{n-1}(u)$ and thus we obtain in case $0 < u < 1$ for the integrand

$$\sum_{n=0}^{\infty} \big| \big(p'_{n-2} - 2p'_{n-1} + p'_n\big)(u) \big| \leq 8.$$

For $u \geq 1$ we obtain for the integrand as an upper bound

$$\begin{aligned} & p_0(u) + p_1(u) \left| 3\frac{1}{u} - 1 \right| + \sum_{n=2}^{\infty} p_n(u) \left| \frac{n(n-1)(n-2)}{u^3} - 3\frac{n(n-1)}{u^2} + 3\frac{n}{u} - 1 \right| \\ & \leq e^{-u}(1+3+u) + \frac{1}{u^3} \sum_{n=2}^{\infty} p_n(u) \left| (n-u)^3 - 3n(n-u) + 2n \right| \\ & \leq e^{-u}(4+u) + \frac{1}{u^3} \left(E|N_u - u|^3 + 3\sqrt{EN_u^2 E(N_u-u)^2} + 2EN_u \right), \end{aligned}$$

where N_u is Poisson with mean u. Note that $E|N_u - u|^m \leq c_m u^{m/2}$ for $m \in \mathbb{N}$ and $u \geq 1$. We also have $EN_u = u$ and $EN_u^2 = u^2 + u$. This yields as an upper bound for the integrand in (2.66) for $u \geq 1$

$$cu^{-\frac{3}{2}} + \frac{1}{u^3} \left(c_3 u^{3/2} + 3\sqrt{(u^2+u) \, c_2 u^1} + 2u \right) \leq cu^{-\frac{3}{2}}.$$

We thus get for E_2

$$\begin{aligned} E_2 & \leq c \, \| f \|_\infty \, \frac{x_k}{t^2} \int_{w^{(k)}}^{w^{'(k)}} \left(u + \frac{1}{2\gamma_k^0} \right)^{-3/2} du \\ & \leq c \, \| f \|_\infty \, \frac{x_k}{t^2} \left| w^{'(k)} - w^{(k)} \right| \left(w^{(k)} + \frac{1}{2\gamma_k^0} \right)^{-3/2} \\ & \leq c \, \| f \|_\infty \, \frac{x_k}{t^2} \frac{h_k}{t} \left(\frac{x_k + t}{2\gamma_k^0 t} \right)^{-3/2} \\ & \leq c \, \| f \|_\infty \, h_k t^{-3/2} \left(t + x_k \right)^{-1/2}. \end{aligned}$$

Together with the bound on E_1 the assertion now follows for $k = i$.

Next investigate the case $k \neq i$. Define

$$\hat{H}_n^1\Big(z, N_t^{'(k)}\Big) = \hat{G}_n\Big(z, N_t^{'(k)}\Big) - \hat{G}_n\Big(0, N_t^{'(k)}\Big),$$
$$\hat{H}_n^2\Big(N_t^{'(k)}, N_t^{(k)}\Big) = \hat{G}_n\Big(0, N_t^{'(k)}\Big) - \hat{G}_n\Big(0, N_t^{(k)}\Big)$$

to get

$$E \leq c\frac{x_i}{t^2}\left|\sum_{n=0}^{\infty} p_n\Big(w^{(i)}\Big)\, E\left[\int \Big(\hat{H}_{n+2}^1 - 2\hat{H}_{n+1}^1 + \hat{H}_n^1\Big)\Big(z, N_t^{'(k)}\Big)\, dQ_h(z)\right]\right|$$
$$+\ c\frac{x_i}{t^2}\left|\sum_{n=0}^{\infty} p_n\Big(w^{(i)}\Big)\, E\left[\int \Big(\hat{H}_{n+2}^2 - 2\hat{H}_{n+1}^2 + \hat{H}_n^2\Big)\Big(N_t^{'(k)}, N_t^{(k)}\Big)\, dQ_h(z)\right]\right|.$$

Recall that the expectation in the definition of $\hat{G}_n\Big(z, N_t^{'(k)}\Big)$ and thus of $\hat{H}_n^1\Big(z, N_t^{'(k)}\Big)$ excludes the random variable $N_t^{'(k)}$. To bound E we thus take expectation w.r.t. $N_t^{'(k)}$, too. Rewriting this yields

$$E \leq c\frac{x_i}{t^2}\sum_{n=0}^{\infty}\left|(p_{n-2} - 2p_{n-1} + p_n)\Big(w^{(i)}\Big)\right|$$
$$\times \sup_{n\geq 0}\left\{E\left[\left|\int \hat{H}_n^1\Big(z, N_t^{'(k)}\Big)\, dQ_h(z)\right|\right] + E\left[\left|\int \hat{H}_n^2\Big(N_t^{'(k)}, N_t^{(k)}\Big)\, dQ_h(z)\right|\right]\right\}$$

and by using $q_n(w) = wp_n(w)$ and $\sum_{n=0}^{\infty} |(q_{n-2} - 2q_{n-1} + q_n)(w)| \leq 2$ again we obtain

$$E \leq c\frac{x_i}{t^2}\frac{1}{w^{(i)}}\sup_{n\geq 0}\left\{E\left[\left|\int \hat{H}_n^1\Big(z, N_t^{'(k)}\Big)\, dQ_h(z)\right|\right]\right.$$
$$\left. + E\left[\left|\int \hat{H}_n^2\Big(N_t^{'(k)}, N_t^{(k)}\Big)\, dQ_h(z)\right|\right]\right\}.$$

Next observe that $\hat{H}_n^1\Big(z, N_t^{'(k)}\Big)$ is zero for $k \in N_2$ and is a first order difference for $k \in N_C$ for which we obtain

$$\left|\int \hat{H}_n^1\Big(z, N_t^{'(k)}\Big)\, dQ_h(z)\right| \leq c\, \| f \|_\infty\, t^{-1}(t + x_k)^{-1}\int z dQ_h(z)$$
$$\leq c\, \| f \|_\infty\, h_k t^{-1/2}(t + x_k)^{-1/2}.$$

The other term can be bounded as follows:

$$\begin{aligned}\left|\hat{H}_n^2\left(N_t^{'(k)}, N_t^{(k)}\right)\right| &\leq \sum_{N=0}^{\infty}\left|p_N\left(w^{'(k)}\right) - p_N\left(w^{(k)}\right)\right| \\ &\qquad \times \left|E\Big[G^{k,i}_{t,x^{N_R}}\Big(x^{N_{C2}}, I_2^{(k)}(t) + R_N^{(k)}, X_0^{'(k)}(t) + S_N^{(k)}, \right. \\ &\qquad\qquad \left. I_2^{(i)}(t) + R_n^{(i)}, X_0^{'(i)}(t) + S_n^{(i)}\Big)\Big]\right| \\ &\leq \sum_{N=0}^{\infty}\left|p_N\left(w^{'(k)}\right) - p_N\left(w^{(k)}\right)\right| \, \| G \|_\infty,\end{aligned}$$

where $w^{(k)} = \frac{x_k}{2\gamma_k^0 t}$ and $w^{'(k)} = \frac{x_k}{2\gamma_k^0 t} + \frac{h_k}{\gamma_k^0 t}$. As done before in the proof of Proposition 2.2.24 we use

$$\sum_{N=0}^{\infty}\left|\int_{w^{(k)}}^{w^{'(k)}} p_N'(u)du\right| \leq ct^{-1/2}h_k(t+x_k)^{-1/2}$$

to finally get with $\| G \|_\infty \leq \| f \|_\infty$

$$\left|\int \hat{H}_n^2\left(N_t^{'(k)}, N_t^{(k)}\right) dQ_h(z)\right| \leq ct^{-1/2}h_k(t+x_k)^{-1/2} \, \| f \|_\infty .$$

Plugging our results into our estimate for E we get

$$E \leq c\frac{x_i}{t^2}\frac{t}{x_i}c \, \| f \|_\infty \, h_k t^{-1/2}(t+x_k)^{-1/2} \leq c \, \| f \|_\infty \, h_k t^{-3/2}(t+x_k)^{-1/2},$$

which proves our assertion. □

Finally we consider the increments in x_k, $k \in N_R$.

Lemma 2.2.30. *If f is a bounded Borel function on $\mathcal{S}_0$, then for all $x, h \in \mathcal{S}_0$, $i \in N_{C2}$ and $k \in N_R$*

$$\left|x_i\frac{\partial^2 P_t f}{\partial x_i^2}(x + h_k e_k) - x_i\frac{\partial^2 P_t f}{\partial x_i^2}(x)\right| \leq \frac{c \, \| f \|_\infty}{t^{3/2}}|h_k| \min_{l \in C_k}\left\{(t+x_l)^{-1/2}\right\}.$$

Proof. Except for the necessary adaptations, already used in the proofs of the preceding assertions, the proof proceeds analogously to Lemma 27 in [7]. □

Continuation of the proof of Proposition 2.2.25. Use Lemmas 2.2.26, 2.2.27 and 2.2.29 in (2.61) together with the calculation following (2.61) to obtain the bound for increments in $x_k, k \in N_{C2}$. Lemma 2.2.30 gives the corresponding bound for increments in $x_k, k \in N_R$ which completes the proof of (2.59). □

2.3 Proof of Uniqueness

As in Section 3, [7], it is relatively straightforward to use the results from the previous sections on the semigroup P_t to prove bounds on the resolvent R_λ of P_t.

We shall then use these bounds to complete the proof of uniqueness of solutions to the martingale problem $MP(\mathcal{A},\nu)$ satisfying Hypothesis 2.1.1 and 2.1.2, where ν is a probability on

$$\mathcal{S} = \left\{ x \in \mathbb{R}_+^d : \prod_{j\in R} \left(\sum_{i\in C_j} x_i + x_j \right) > 0 \right\}$$

(recall (2.3) and Lemma 2.1.5) and

$$\mathcal{A}f(x) = \sum_{j\in R} \gamma_j(x) \left(\sum_{i\in C_j} x_i \right) x_j f_{jj}(x) + \sum_{j\notin R} \gamma_j(x) x_j f_{jj}(x) + \sum_{j\in V} b_j(x) f_j(x). \tag{2.67}$$

The proof of uniqueness is identical to the one in [7] except for minor changes such as the replacement of x_{c_j} by $\sum_{i\in C_j} x_i$ at the appropriate places. Note in particular the change in the definition of the state space $\mathcal{S}$.

In what follows we shall give a sketch of the proofs and indicate where statements have to be modified. For explicit calculations the reader is referred to [7], Sections 3 and 4.

Notation 2.3.1. For $i \in N_{C2}$ let

$$\bar{y}_i = \left(\{y_j\}_{j\in\bar{R}_i}, y_i \right), \bar{y}_i \bar{e}_i = \sum_{j\in\bar{R}_i} y_j e_j + y_i e_i \text{ and } \bar{\mathbb{R}}_i = \mathbb{R}^{|\bar{R}_i|} \times \mathbb{R}_+, \tag{2.68}$$

where we understand this to be $\bar{y}_i = (y_i)$ in case $i \in N_2$, i.e. $\bar{R}_i = \emptyset$. For $f \in \mathcal{C}_b^2(\mathcal{S}_0)$ let

$$\frac{\partial f}{\partial \bar{x}_i} = \left(\left\{ \frac{\partial}{\partial x_j} f \right\}_{j\in\bar{R}_i}, \frac{\partial}{\partial x_i} f \right), \left| \frac{\partial f}{\partial \bar{x}_i} \right| = \sum_{j\in\bar{R}_i} \left| \frac{\partial}{\partial x_j} f \right| + \left| \frac{\partial}{\partial x_i} f \right| \tag{2.69}$$

and

$$\left\| \left| \frac{\partial f}{\partial \bar{x}_i} \right| \right\|_\infty = \sup \left\{ \left| \frac{\partial f}{\partial \bar{x}_i}(x) \right| : x \in \mathcal{S}_0 \right\}, \tag{2.70}$$

where $\mathcal{S}_0 = \{x \in \mathbb{R}^d : x_i \geq 0 \text{ for all } i \in N_{C2}\}$ as defined in (2.9). Also introduce

$$\Delta_i f = \left(\left\{ x_i \frac{\partial^2}{\partial x_j^2} f \right\}_{j\in\bar{R}_i}, x_i \frac{\partial^2}{\partial x_i^2} f \right).$$

Define $|\Delta_i f|$ and $\| \Delta_i f \|_\infty$ similarly to (2.69) and (2.70).

With the help of these notations $\mathcal{A}^0$ (see (2.6)) can be rewritten to

$$\begin{aligned}\mathcal{A}^0 f(x) &= \sum_{j\in V} b_j^0 f_j(x) + \sum_{j\in N_R} \gamma_j^0 \left(\sum_{i\in C_j} x_i\right) f_{jj}(x) + \sum_{i\in N_{C2}} \gamma_i^0 x_i f_{ii}(x) \qquad (2.71)\\ &= \sum_{i\in N_{C2}} \left\langle \underline{b^0}_i, \frac{\partial f}{\partial \bar{x}_i}(x)\right\rangle + \sum_{i\in N_{C2}} \left\langle \underline{\gamma^0}_i, \Delta_i f(x)\right\rangle,\end{aligned}$$

where $\langle\cdot,\cdot\rangle$ denotes the standard scalar product in $\mathbb{R}^k, k\in\mathbb{N}$. To prevent over-counting in case $\bar{R}_{i_1}\cap\bar{R}_{i_2}\neq\emptyset$ for $i_1\neq i_2$, $i_1,i_2\in N_C$ (see also definition (2.68)) the vector $\overline{b^0}_i$ was replaced by $\underline{b^0}_i$ in the above formula, where $\underline{b^0}_i$ has certain coordinates set to zero so that the above equality holds. The same applies to the vector $\underline{\gamma^0}_i$. The details are left to the interested reader.

Theorem 2.3.2. *There is a constant c such that for all $f\in\mathcal{C}_w^\alpha(\mathcal{S}_0), \lambda\geq 1$ and $k\in N_{C2}$,*

$$(a)\quad \left|\left|\frac{\partial R_\lambda f}{\partial \bar{x}_k}\right|\right|_\infty + ||\Delta_k R_\lambda f||_\infty \leq c\lambda^{-\alpha/2}|f|_{\mathcal{C}_w^\alpha}.$$

$$(b)\quad \left|\frac{\partial R_\lambda f}{\partial \bar{x}_k}\right|_{\mathcal{C}_w^\alpha} + |\Delta_k R_\lambda f|_{\mathcal{C}_w^\alpha} \leq c|f|_{\mathcal{C}_w^\alpha}.$$

Note. This result is slightly weaker than the corresponding Theorem 34 in [7] as $|f|_{\alpha,k}$ is replaced by $|f|_{\mathcal{C}_w^\alpha}$ in (a).

Proof. Firstly we obtain a result similar to Proposition 30 in [7]. This is an easy consequence of Proposition 2.2.13 and Proposition 2.2.16, using the equivalence of norms shown in Theorem 2.2.20 and states that there is a constant c such that

(a) For all $f\in\mathcal{C}_w^\alpha(\mathcal{S}_0), t>0, x\in\mathcal{S}_0$, and $i\in N_{C2}$,

$$\left|\frac{\partial P_t f}{\partial \bar{x}_i}(x)\right| \leq c|f|_{\mathcal{C}_w^\alpha} t^{\alpha/2-1/2}(t+x_i)^{-1/2} \leq c|f|_{\mathcal{C}_w^\alpha} t^{\alpha/2-1}, \qquad (2.72)$$

and

$$\|\Delta_i P_t f\|_\infty \leq c|f|_{\mathcal{C}_w^\alpha} t^{\alpha/2-1}. \qquad (2.73)$$

(b) For all f bounded and Borel on $\mathcal{S}_0$ and all $i\in N_{C2}$,

$$\left|\left|\frac{\partial P_t f}{\partial \bar{x}_i}\right|\right|_\infty \leq c\,\|f\|_\infty\, t^{-1}.$$

Note in particular that Theorem 2.2.20 gave $\mathcal{C}_w^\alpha=\mathcal{S}^\alpha$ and that every function in $\mathcal{C}_w^\alpha(\mathcal{S}_0)$ is by definition bounded.

Secondly, an easy consequence of Propositions 2.2.24, 2.2.25 and the triangle inequality, using the equivalence of norms shown in Theorem 2.2.20 and the equivalence of the maximum norm and Euclidean norm of finite dimensional vectors, is a result similar to Proposition 32, [7]: There is a constant c such that for all $f \in \mathcal{C}_w^\alpha(\mathcal{S}_0), i, k \in N_{C2}$ and $\bar{h}_i \in \bar{\mathbb{R}}_i$,

(a)

$$\left| \frac{\partial P_t f}{\partial \bar{x}_k}\big(x + \bar{h}_i \bar{e}_i\big) - \frac{\partial P_t f}{\partial \bar{x}_k}(x) \right| \leq c|f|_{\mathcal{C}_w^\alpha} t^{-3/2+\alpha/2}(t + x_i)^{-1/2} \left|\bar{h}_i\right|, \qquad (2.74)$$

(b)

$$\left|\Delta_k(P_t f)\big(x + \bar{h}_i \bar{e}_i\big) - \Delta_k(P_t f)(x)\right| \leq c|f|_{\mathcal{C}_w^\alpha} t^{-3/2+\alpha/2}(t + x_i)^{-1/2} \left|\bar{h}_i\right|. \quad (2.75)$$

Finally recall that $R_\lambda f(x) = \int_0^\infty e^{-\lambda t} P_t f(x) dt$ is the resolvent associated with P_t. Now the remainder of the proof works as in the proof of Theorem 34 in [7]: Part (a) of Theorem 2.3.2 is obtained by integrating (2.72) resp. (2.73) over time. Part (b) follows by integrating (2.72) resp. (2.73) over the time-interval from zero to some fixed value $\tilde{t} > 0$ and (2.74) resp. (2.75) over the time interval from $\tilde{t}$ to infinity. Appropriate choices for $\tilde{t}$ now yield the required bounds. Here the choices of $\tilde{t}$ are in fact easier due to the replacement of $|\cdot|_{\alpha,i}$ in [7] by $|\cdot|_{\mathcal{C}_w^\alpha}$. □

Proof of Theorem 2.1.6. The existence of a solution to the martingale problem for $MP(\mathcal{A}, \nu)$ follows by standard methods (a result of Skorokhod yields existence of approximating solutions, then use a tightness-argument), e.g. see the proof of Theorem 1.1 in [1]. Note in particular that Lemma 2.1.5 ensures that solutions remain in $\mathcal{S} \subset \mathbb{R}_+^d$. The uniform boundedness in M of the term $E\Big[\sum_i \Big|X_T^{M,i}\Big|\Big]$ that appears in the proof of Theorem 1.1 in [1] can easily be replaced by the uniform boundedness in M of $E\Big[\sum_{i \in V} (X_T^{M,i})^2\Big]$ via a Gronwall-type argument.

At the end of this section we shall reduce the proof of uniqueness to the following theorem. The theorem investigates uniqueness of a perturbation of the operator $\mathcal{A}^0$ as defined in (2.6) (also refer to (2.71)) with coefficients satisfying (2.7) and (2.8). $\mathcal{A}^0$ is the generator of a unique diffusion on $\mathcal{S}\big(x^0\big)$ given by (2.9) with semigroup P_t and resolvent R_λ given by (2.11). For the definition of M^0 refer to (2.14).

In what follows $x^0 \in \mathcal{S}$ will be arbitrarily fixed.

Theorem 2.3.3. *Assume that*

$$\begin{aligned} \tilde{\mathcal{A}} f(x) = & \sum_{j \in N_R} \tilde{\gamma}_j(x) \left(\sum_{i \in C_j} x_i \right) f_{jj}(x) \\ & + \sum_{j \in N_{C2}} \tilde{\gamma}_j(x) x_j f_{jj}(x) + \sum_{j \in V} \tilde{b}_j(x) f_j(x), \ x \in \mathcal{S}\big(x^0\big), \end{aligned} \qquad (2.76)$$

where $\tilde{b}_k : \mathcal{S}\big(x^0\big) \to \mathbb{R}$ *and* $\tilde{\gamma}_k : \mathcal{S}\big(x^0\big) \to (0, \infty)$,

$$\tilde{\Gamma} = \sum_{k=1}^d ||\tilde{\gamma}_k||_{\mathcal{C}_w^\alpha} + \left|\left|\tilde{b}_k\right|\right|_{\mathcal{C}_w^\alpha} < \infty.$$

Let

$$\tilde{\epsilon}_0 = \sum_{k=1}^{d} \left|\left|\tilde{\gamma}_k - \gamma_k^0\right|\right|_\infty + \left|\left|\tilde{b}_k - b_k^0\right|\right|_\infty ,$$

where $b_k^0, \gamma_k^0, k \in V$ *satisfy (2.7). Let* $\mathcal{B}f = (\tilde{\mathcal{A}} - \mathcal{A}^0)f$.

(a) There exists $\epsilon_1 = \epsilon_1(M^0) > 0$ *and* $\lambda_1 = \lambda_1(M^0, \tilde{\Gamma}) \geq 0$ *such that if* $\tilde{\epsilon}_0 \leq \epsilon_1$ *and* $\lambda \geq \lambda_1$ *then* $\mathcal{B}R_\lambda : \mathcal{C}_w^\alpha \to \mathcal{C}_w^\alpha$ *is a bounded operator with* $\| \mathcal{B}R_\lambda \| \leq 1/2$.

(b) If we assume additionally that $\tilde{\gamma}_k$ *and* $\tilde{b}_k$ *are Hölder continuous of index* $\alpha \in (0,1)$, *constant outside a compact set and* $\tilde{b}_k|_{\{x_k=0\}} \geq 0$ *for all* $k \in V \backslash N_R$, *then the martingale problem MP(*$\tilde{\mathcal{A}}, \nu$*) has a unique solution for each probability* ν *on* $\mathcal{S}(x^0)$.

Proof. Let $\tilde{R}_\lambda$ be the associated resolvent operator of the perturbation operator $\tilde{\mathcal{A}}$. Using the definition $\mathcal{B} = \tilde{\mathcal{A}} - \mathcal{A}^0$ and recalling (2.71) we get for $f \in \mathcal{C}_w^\alpha$ that

$$\begin{aligned} \| \mathcal{B}R_\lambda f \|_{\mathcal{C}_w^\alpha} \leq & \sum_{i \in N_{C2}} \left|\left| \left\langle \underline{\left(\tilde{b}(x) - b^0\right)}_i, \frac{\partial R_\lambda f}{\partial \bar{x}_i}(x) \right\rangle \right|\right|_{\mathcal{C}_w^\alpha} \\ & + \sum_{i \in N_{C2}} \left|\left| \left\langle \underline{\left(\tilde{\gamma}(x) - \gamma^0\right)}_i, \Delta_i R_\lambda f(x) \right\rangle \right|\right|_{\mathcal{C}_w^\alpha} . \end{aligned}$$

Using (2.46) (recall in particular the discussion on the reasons for using two different norms from Remark 2.2.21) we obtain for instance for arbitrary $i \in N_C$ and $j \in \bar{R}_i$

$$\begin{aligned} & \left| \left(\tilde{b}_j(x) - b_j^0\right) \frac{\partial R_\lambda f}{\partial x_j}(x) \right|_{\mathcal{C}_w^\alpha} \\ & \leq c \left[\left|\left| \tilde{b}_j(x) - b_j^0 \right|\right|_{\mathcal{C}_w^\alpha} \left|\left| \frac{\partial R_\lambda f}{\partial x_j}(x) \right|\right|_\infty + \left|\left| \tilde{b}_j(x) - b_j^0 \right|\right|_\infty \left| \frac{\partial R_\lambda f}{\partial x_j}(x) \right|_\alpha \right] \\ & \leq c \left[\left(\tilde{\Gamma} + M^0\right) \lambda^{-\alpha/2} |f|_{\mathcal{C}_w^\alpha} + \tilde{\epsilon}_0 |f|_{\mathcal{C}_w^\alpha} \right] \end{aligned}$$

by Theorem 2.3.2, (2.50) and the assumptions of this theorem. By arguing similarly for the other terms we get indeed $\| \mathcal{B}R_\lambda f \|_{\mathcal{C}_w^\alpha} \leq \frac{1}{2} \| f \|_{\mathcal{C}_w^\alpha}$ for λ big enough thus finishing the proof of part (a).

For part (b) we proceed as in the proof of [7], Theorem 37. The proof of Theorem 37 in [7] involves the proof of Lemma 38 in [7], where one shows that for $f \in \mathcal{C}_w^\alpha$

$$\tilde{R}_\lambda f = R_\lambda f + \tilde{R}_\lambda \mathcal{B} R_\lambda f. \tag{2.77}$$

Note that the proof of Lemma 38 relies amongst others on an estimate, derived in Corollary 33 of [7], which we now obtain for free in Proposition 2.2.11 as we treated all vertices in one step only.

The proof of Theorem 37 now concludes as follows. Iteration of (2.77) yields

$$\tilde{R}_\lambda f(x) = \sum_{n=0}^{\infty} R_\lambda((\mathcal{B}R_\lambda)^n f)(x).$$

Using $\| \mathcal{B}R_\lambda \|_{\mathcal{C}_w^\alpha} \leq 1/2$ from part (a) and $\| f \|_\infty \leq \| f \|_{\mathcal{C}_w^\alpha}$ we get

$$\lambda \| R_\lambda((\mathcal{B}R_\lambda)^n f) \|_\infty \leq \| (\mathcal{B}R_\lambda)^n f \|_\infty \leq \| (\mathcal{B}R_\lambda)^n f \|_{\mathcal{C}_w^\alpha} \leq 2^{-n} \| f \|_{\mathcal{C}_w^\alpha} .$$

Thus the series converges uniformly and the error term approaches zero. The uniqueness of $MP(\tilde{\mathcal{A}}, \nu)$ now follows from the uniqueness of its resolvents $\tilde{R}_\lambda$. ☐

Continuation of the proof of Theorem 2.1.6. Recall "Step 1: Reduction of the problem", in Subsection 2.1.5. The remainder of the proof of uniqueness of $MP(\mathcal{A}, \delta_{x^0})$ works analogously to [7] (compare the proof of Theorem 4 on pp. 380-382 in [7]) except for minor changes, making again use of Lemma 2.1.5. The main step consists in using a localization argument of [13] (see e.g. the argument in the proof of Theorem 1.2 of [4]), which basically states that it is enough if for each $x^0 \in \mathcal{S}$ the martingale problem $MP(\tilde{\mathcal{A}}, \delta_{x^0})$ has a unique solution, where $b_i = \tilde{b}_i$ and $\gamma_i = \tilde{\gamma}_i$ agree on some neighbourhood of x^0. By comparing the definition of $\mathcal{A}$ (see (2.67)) and $\tilde{\mathcal{A}}$ (see (2.76)) one chooses

$$\begin{aligned}
\tilde{b}_k(x) &= b_k(x) \text{ for all } k \in V, \\
\tilde{\gamma}_j(x) &= x_j \gamma_j(x) \text{ for } j \in N_R, \\
\tilde{\gamma}_j(x) &= \left(\sum_{i \in C_j} x_i \right) \gamma_j(x) \text{ for } j \in R \backslash N_R \\
\tilde{\gamma}_j(x) &= \gamma_j(x) \text{ for } j \notin R.
\end{aligned}$$

By setting

$$b_k^0 \equiv \tilde{b}_k(x^0) \text{ and } \gamma_k^0 \equiv \tilde{\gamma}_k(x^0)$$

and choosing $\tilde{b}_k$ and $\tilde{\gamma}_k$ in appropriate ways, the assumptions of Theorem 2.3.3(a), (b) will be satisfied in case $b_k^0 \geq 0$ for all $k \in N_2$ (and hence by Hypothesis 2.1.2 for all $k \in N_{C2}$). In particular the boundedness and continuity of the coefficients of $\tilde{\mathcal{A}}$ will allow us to choose $\tilde{\epsilon}_0$ arbitrarily small. In case there exists $k \in N_2$ such that $b_k^0 < 0$ a Girsanov argument as in the proof of Theorem 1.2 of [4] allows the reduction of the latter case to the former case. ☐

Bibliography

[1] Athreya, S.R. and Barlow, M.T. and Bass, R.F. and Perkins, E.A. Degenerate stochastic differential equations and super-Markov chains. *Probab. Theory Related Fields* (2002) **123**, 484–520. MR1921011

[2] Athreya, S.R. and Bass, R.F. and Perkins, E.A. Hölder norm estimates for elliptic operators on finite and infinite-dimensional spaces. *Trans. Amer. Math. Soc.* (2005) **357**, 5001–5029 (electronic). MR2165395

[3] Bass, R.F. *Diffusions and Elliptic Operators*. Springer, New York, 1998. MR1483890

[4] Bass, R.F. and Perkins, E.A. Degenerate stochastic differential equations with Hölder continuous coefficients and super-Markov chains. *Trans. Amer. Math. Soc.* (2003) **355**, 373–405 (electronic). MR1928092

[5] Bass, R.F. and Perkins, E.A. Degenerate stochastic differential equations arising from catalytic branching networks. *Electron. J. Probab.* (2008) **13**, 1808–1885. MR2448130

[6] Dawson, D.A. and Greven, A. and den Hollander, F. and Sun, R. and Swart, J.M. The renormalization transformation for two-type branching models. *Ann. Inst. H. Poincaré Probab. Statist.* (2008) **44**, 1038–1077. MR2469334

[7] Dawson, D.A. and Perkins, E.A. On the uniqueness problem for catalytic branching networks and other singular diffusions. *Illinois J. Math.* (2006) **50**, 323–383 (electronic). MR2247832

[8] Eigen, M. and Schuster, P. *The Hypercycle: a principle of natural self-organization*. Springer, Berlin, 1979.

[9] Hofbauer, J. and Sigmund, K. *The Theory of Evolution and Dynamical Systems*. London Math. Soc. Stud. Texts, vol. 7, Cambridge Univ. Press, Cambridge, 1988. MR1071180

[10] Mytnik, L. Uniqueness for a mutually catalytic branching model. *Probab. Theory Related Fields* (1998) **112**, 245–253. MR1653845

[11] Perkins, E.A. *Dawson-Watanabe superprocesses and measure-valued diffusions*. Lectures on Probability Theory and Statistics (Saint-Flour, 1999), 125–324, Lecture Notes in Math., 1781, Springer, Berlin, 2002. MR1915445

[12] Rogers, L.C.G. and Williams, D. *Diffusions, Markov Processes, and Martingales, vol. 2*, Reprint of the second (1994) edition. Cambridge Mathematical Univ. Press, Cambridge, 2000. MR1780932

[13] Stroock, D.W. and Varadhan, S.R.S. *Multidimensional Diffusion Processes*. Grundlehren Math. Wiss., vol. 233, Springer, Berlin-New York, 1979. MR532498

Chapter 3

Long-term Behaviour of a Cyclic Catalytic Branching System[1]

3.1 Introduction

3.1.1 Basics

In this paper we investigate the long-term behaviour of the following system of stochastic differential equations (SDEs) for $d \geq 2$:

$$dX_t^i = \sqrt{2\gamma^i X_t^i X_t^{i+1}} dB_t^i + \sum_{j=1}^{d} X_t^j q_{ji} dt, \; i \in \{1, \ldots, d\}, \tag{3.1}$$

where $X_t^{d+1} \equiv X_t^1$. We shall assume the γ^i and $q_{ji}, i \neq j$ to be given positive constants and the $X_0^i \geq 0, i \in \{1, \ldots, d\}$ to be given initial conditions. (q_{ji}) is a Q-matrix modelling mutations from type j to type i.

This system involves both cyclic catalytic branching and mutation between types. The extension of the cyclic setup to arbitrary networks (see Subsection 3.2.6 at the end of this paper) is straightforward. Existence of solutions shall be shown by standard methods. To show weak uniqueness we shall employ the results of Dawson and Perkins [3] once we show that a solution does not hit $0 \in \mathbb{R}^d$ in finite time.

The given system of SDEs can be understood as a stochastic analogue to a system of ODEs for the concentrations $y_j, j = 1, \ldots, d$ of a type T_j. Then $y_j / \dot{y}_j$ corresponds to the rate of growth of type T_j and one obtains the following ODEs (see Hofbauer and Sigmund [6]): for independent replication $\dot{y}_j = b_j y_j$, autocatalytic replication $\dot{y}_j = \gamma_j y_j^2$ and catalytic replication $\dot{y}_j = \gamma_j \big(\sum_{i \in C_j} y_i\big) y_j$. In the cyclic catalytic case type T_{j+1} catalyzes the replication of type j, i.e. the growth of type j is proportional to the mass of type $j+1$ present at time t. The cyclic catalytic case represents the simplest form of mutual help between different types. It was firstly introduced by Eigen and Schuster (see Eigen and Schuster [4]).

[1]A version of this chapter will be submitted for publication. Kliem, S.M. (2009) Long-term Behaviour of a Cyclic Catalytic Branching System.

The system of SDEs can be obtained as a limit of branching particle systems. The growth rate of types in the ODE setting now corresponds to the branching rate in the stochastic setting, i.e. type j branches at a rate proportional to the mass of type $j+1$ at time t.

Results on weak uniqueness for catalytic branching networks can be found for instance in [3] and Kliem [9]. The former proved weak uniqueness for catalytic replication under the restriction to networks with at most one catalyst per reactant, which includes the hypercyclic case. The latter removed this restriction. Both papers allow more general diffusion- and drift- coefficients under some Hölder-continuity conditions. These conditions were weakened in Bass and Perkins [1] to continuity only.

Our main interest shall be the long-time behaviour of the above system. In particular, we shall investigate survival and coexistence of types. Such questions naturally arise in biological competition models. For instance, Fleischmann and Xiong [5] investigated a cyclically catalytic super-Brownian motion. They showed global segregation (noncoexistence) of neighbouring types in the limit and other results on the finite time survival-extinction but they were not able to determine, if the overall sum dies out in the limit or not.

In this paper we shall show that in our SDE-setup the overall sum converges to zero but does not hit zero in finite time. To further analyze the relative behaviour of types while they approach zero, we turned our attention to the normalized processes $Y_t^i \equiv X_t^i / \sum_j X_t^j$ - note that $X_t^i/X_t^j = Y_t^i/Y_t^j$ - and showed weak convergence to a unique stationary distribution that does not charge the set where at least one of the coordinates is zero.

3.1.2 Main results and outline of the paper

As a first step we shall show existence and nonnegativity of solutions $X_t^i, i \in \{1, \ldots, d\}$ to the above SDE by standard methods in Subsection 3.2.1. As a next step we shall prove in Subsection 3.2.2 that the sum of all coordinates, i.e. $S_t \equiv \sum_{i=1}^d X_t^i$, converges to zero but does not hit zero in finite time a.s. We then establish the weak uniqueness of the system by Theorem 4 of [3] or Theorem 1.6 of [9].

Secondly, from Subsection 3.2.3 on we shall change our focus to the normalized processes, i.e. to $Y_t^i = X_t^i/S_t$ to get some insight on the relative behaviour of types. Existence of solutions follows again by standard methods and the weak uniqueness of solutions in $[0,1]^d$ follows by establishing a connection between the system at hand and the original system of SDEs. In Subsection 3.2.4 we show that any stationary distribution for Y_t does not charge the set where at least one of the coordinate processes becomes extinct. We shall use this result in Subsection 3.2.5 to prove weak convergence to a unique stationary distribution by adapting the proof of Theorem 2.3 of Dawson, Greven, den Hollander, Sun and Swart [2] to our setup.

Finally, in Subsection 3.2.7 we shall give a complete analysis of the case $d = 2$ by using methods of speed and scale.

3.2 Main Results

3.2.1 Existence and nonnegativity

Let $(\Omega, \mathcal{F}, (\mathcal{F})_t, \mathbb{P})$ be a filtered probability space that satisfies the usual conditions (cf. Rogers and Williams [10], Introduction to Chapter IV). Consider the following system of SDEs for $d \geq 2$:

$$dX_t^i = \sqrt{2\gamma^i X_t^i X_t^{i+1}} dB_t^i + \sum_{j=1}^{d} X_t^j q_{ji} dt, \ i \in \{1, \dots, d\}, \tag{3.2}$$

where $X_t^{d+1} \equiv X_t^1$. We shall assume the γ^i and $q_{ji}, i \neq j$ to be given strictly positive constants and the $X_0^i \geq 0, i \in \{1, \dots, d\}$ to be given initial conditions. As the q_{ji} model mutations from type j to type i we impose

$$q_{ii} = -\sum_{j:j \neq i} q_{ij} \iff \sum_j q_{ij} = 0. \tag{3.3}$$

Let

$$q_{max} = \max_{1 \leq i,j \leq d} |q_{ij}|, \ q_{min} = \min_{1 \leq i,j \leq d} |q_{ij}| > 0 \text{ and } \gamma_{max} = \max_{1 \leq i \leq d} \gamma_i > 0.$$

First we shall investigate the existence of solutions to (3.2).

Lemma 3.2.1. *There exists a solution to the given system of SDEs (3.2).*

Reference for the proof. Existence follows by standard methods, see for instance Theorem V.3.10 in Ethier and Kurtz [7]. □

Next, we shall show that all solutions to (3.2) stay in the first quadrant (here we replaced the terms under the square root with their absolute values to be able to consider solutions on all of $\mathbb{R}^d$). For this purpose we shall first show that the local time of the coordinate processes at zero is zero.

Corollary 3.2.2. *Let $i \in \{1, \dots, d\}$ be arbitrarily fixed. Then the local time l_t^0 at zero of the process X_t^i is zero.*

Proof. The proof proceeds along the lines of standard techniques for local times. Let $i \in \{1, \dots, d\}$ be arbitrarily fixed. By [10], IV.(45.3) ("occupation density formula") we have for $\varphi(x) = 1_{[0,\epsilon]}(x)$, $\epsilon > 0$,

$$\int_0^t \varphi(X_s^i) d < X^i >_s = \int_0^t 1_{\{0 \leq X_s^i \leq \epsilon\}} 2\gamma^i \left| X_s^i X_s^{i+1} \right| ds = \int_0^\epsilon l_t^a da = \int_{\mathbb{R}} l_t^a \varphi(a) da.$$

Next recall from Theorem IV.(44.2) that without loss of generality l_t^a is right-continuous in a. We also know that the processes under consideration are continuous. Hence,

$$0 \leq l_t^0 = \lim_{\epsilon \downarrow 0^+} \frac{1}{\epsilon} \int_0^\epsilon l_t^a da \leq \lim_{\epsilon \downarrow 0^+} \frac{1}{\epsilon} \int_0^t 1_{\{0 < X_s^i \leq \epsilon\}} 2\gamma^i \epsilon \left| X_s^{i+1} \right| ds = 0,$$

the last by dominated convergence, proving the assumption. □

Notation 3.2.3. In what follows we shall denote the martingale part of X_t^i by

$$\mathcal{M}_t^i \equiv \int_0^t \sqrt{2\gamma^i X_s^i X_s^{i+1}} dB_s^i, \ i \in \{1, \ldots, d\}.$$

Lemma 3.2.4. *The processes X_t^i, $i \in \{1, \ldots, d\}$ are nonnegative if we start at $X_0^i \geq 0, \forall i \in \{1, \ldots, d\}$. We also obtain that the $\mathcal{M}_t^i, i \in \{1, \ldots, d\}$ are martingales.*

Proof. To the purpose of proving this Lemma we shall use that

$$\left(X_t^i\right)^- = \int_0^t -1_{\{X_s^i \leq 0\}} dX_s^i + \frac{1}{2} l_t^0 = \int_0^t -1_{\{X_s^i \leq 0\}} dX_s^i, \tag{3.4}$$

the last by Corollary 3.2.2

Before we continue, observe that $\mathcal{M}_t^i \equiv \int_0^t \sqrt{2\gamma^i X_s^i X_s^{i+1}} dB_s^i$ is a martingale. Indeed, $\mathcal{M}^i$ is a continuous local martingale and so it suffices to show that $\mathbb{E}\left[< \mathcal{M}^i >_t\right] < \infty$ for all $t > 0$.

To show this defines a sequence of stopping times $T_n \equiv \inf\{t \geq 0 : \max_{i=1}^d |X_t^i| \geq n\}$. As Cauchy Schwarz' inequality yields

$$\mathbb{E}\left[< \mathcal{M}^i >_{t \wedge T_n}\right] \leq C \int_0^t \sqrt{\mathbb{E}\left[(X_{s \wedge T_n}^i)^2\right] \mathbb{E}\left[(X_{s \wedge T_n}^{i+1})^2\right]} ds < \infty, \tag{3.5}$$

$\mathcal{M}_{\cdot \wedge T_n}^i$ is a continuous martingale. In particular, $\mathbb{E}\left[\left(\mathcal{M}_{t \wedge T_n}^i\right)^2\right] = \mathbb{E}\left[< \mathcal{M}_{\cdot \wedge T_n}^i >_t\right]$ and we obtain that

$$\begin{aligned}\mathbb{E}\left[(X_{t \wedge T_n}^i)^2\right] &\leq C \left\{ (X_0^i)^2 + \mathbb{E}\left[\left(\mathcal{M}_{t \wedge T_n}^i\right)^2\right] + \mathbb{E}\left[\left(\int_0^{t \wedge T_n} \sum_{j=1}^d X_s^j q_{ji} ds\right)^2\right] \right\} \\ &\stackrel{(3.5)}{\leq} C \left(1 + \int_0^t \max_{i=1}^d \mathbb{E}\left[(X_{s \wedge T_n}^i)^2\right] ds\right).\end{aligned}$$

Hence Gronwall's lemma gives $\max_{i=1}^d \mathbb{E}\left[(X_{t \wedge T_n}^i)^2\right] \leq Ce^{Ct}$. As $T_n \to \infty$ for $n \to \infty$ we can now apply the monotone convergence theorem in (3.5) to get $\mathbb{E}\left[< \mathcal{M}^i >_t\right] < \infty$ for all $t > 0$. Thus $\mathcal{M}^i$ is indeed a continuous martingale.

Taking expectations in (3.4) this implies

$$\mathbb{E}\left[\left(X_t^i\right)^-\right] \leq \mathbb{E}\left[\int_0^t \sum_{j=1}^d 1_{X_s^j \leq 0}(-X_s^j) q_{ji} ds\right].$$

Sum both sides over i and use (3.3) to obtain $0 \leq \sum_i \mathbb{E}\left[\left(X_t^i\right)^-\right] \leq 0$. Thus $X_t^i \geq 0$ a.s. for all $i \in \{1, \ldots, d\}$ and $t \geq 0$. □

3.2.2 The overall sum and uniqueness

In what follows we shall investigate the behaviour of our system for $t \to \infty$. We shall show that the sum of all coordinates converges to zero but does not hit zero at a finite time. At the end of this Subsection we shall use this result to establish the weak uniqueness of solutions to (3.2).

Notation 3.2.5. Let $S_t \equiv \sum_{i=1}^d X_t^i$.

Corollary 3.2.6. *S_t converges a.s. for $t \to \infty$.*

Proof. First note that by Lemma 3.2.4, S_t is a nonnegative process. Using (3.2) we obtain for S_t

$$dS_t = \sum_{i=1}^d \sqrt{2\gamma^i X_t^i X_t^{i+1}} dB_t^i + \sum_{i=1}^d \sum_{j=1}^d X_t^j q_{ji} dt \stackrel{(3.3)}{=} \sum_{i=1}^d \sqrt{2\gamma^i X_t^i X_t^{i+1}} dB_t^i.$$

Using that the $\mathcal{M}_t^i$ are martingales as shown in Lemma 3.2.4, we obtain that S_t is a nonnegative martingale and thus a.s. convergent. □

Lemma 3.2.7. *$S_t > 0$ for all $t \geq 0$ a.s., given that $S_0 > 0$.*

Proof. First observe that

$$< S >_t = \int_0^t \sum_{i=1}^d 2\gamma^i X_s^i X_s^{i+1} ds \leq 2\gamma_{max} \int_0^t S_s^2 ds.$$

Next we shall use a time-change to be able to take advantage of this inequality. Let

$$I_s \equiv \sum_{i=1}^d 2\gamma^i X_s^i X_s^{i+1} \text{ and } C_t \equiv \int_0^t \frac{I_s}{2\gamma_{max} S_s^2} ds \text{ for } t \leq \tau,$$

where $\tau \equiv \inf\{t \geq 0 : \exists \epsilon > 0 \text{ such that } I_s = 0 \ \forall s \in [t, t+\epsilon]\}$.

Note that if $S_{t_0} = 0$ for some $t_0 > 0$, then $S_t = 0$ for all $t \geq t_0$ by the optional sampling theorem as S_t is a continuous nonnegative martingale. As $I_s \leq 2\gamma_{max} S_s^2$ we therefore have $S_s > 0$ for all $s < \tau$.

Also note that $C_t < \infty$ for all $t < \tau$ as $0 \leq I_s$ and $I_s \leq 2\gamma_{max} S_s^2$ which yields

$$0 \leq \frac{I_s}{2\gamma_{max} S_s^2} \leq 1 \text{ for all } 0 \leq s < \tau. \tag{3.6}$$

In particular the definition of τ implies that C_t is a continuous strictly increasing function defining a homeomorphism between $[0, \tau]$ and $[0, \xi]$ for $\xi < \infty$ respectively $[0, \infty)$ if $\xi = \infty$ (let us also denote this by $[0, \xi]$), where $\xi \equiv C_\tau$.

Let $D : [0, \xi] \to [0, \tau]$ be the continuous strictly increasing inverse to C_t.

Under this time-change we now obtain for S_t,

$$Z_t \equiv S_{D_t}, t \leq \xi,$$

where

$$< Z >_t = < S >_{D_t} = \int_0^{D_t} I_s ds = \int_0^t I_{D_s} \frac{2\gamma_{max} S_{D_s}^2}{I_{D_s}} ds = 2\gamma_{max} \int_0^t Z_s^2 ds,$$

i.e.

$$d < Z >_t = 2\gamma_{max} Z_t^2 dt, \ \forall t \leq \xi.$$

Thus Z is a geometric Brownian motion (see for instance Karatzas and Shreve [8], Exercise 5.5.31). In particular $Z_t > 0 \ \forall t \leq \xi$ if $\xi < \infty$ respectively $Z_t > 0$ $\forall t < \xi$ if $\xi = \infty$ follows, given that $Z_0 = S_0 > 0$. We therefore obtain

$$\begin{cases} S_{D_t} > 0 \ \forall t \leq \xi \Rightarrow S_t > 0 \ \forall t \leq \tau, & \text{if } \xi < \infty \text{ respectively} \\ S_{D_t} > 0 \ \forall t < \xi \Rightarrow S_t > 0 \ \forall t < \tau, & \text{if } \xi = \infty \ \left(\overset{(3.6)}{\Rightarrow} \tau = \infty \right). \end{cases} \tag{3.7}$$

In what follows let $\eta = \inf\{t \geq 0 : S_t = 0\}$. To finish our proof it remains to show that $\eta = \infty$ a.s. By (3.7) it can easily be seen that we have $\tau \leq \eta$ and by continuity of S_t the infimum in the definition of η is attained given $\eta < \infty$. Finally note that $S_T = 0$ implies $S_t = 0 \ \forall t \geq T$ as S_t is a nonnegative martingale. Indeed, T is a stopping time and so this follows from the optional sampling theorem (see for instance [7], II.2.13).

Let us suppose by contradiction that $\eta < \infty$. We are left with two cases.

1. case. $\tau(\omega) = \eta(\omega) < \infty$.

This yields a contradiction as $S_\tau = S_\eta = 0$ by definition of η but $S_\tau > 0$ by (3.7).

Before investigating the second case we shall need a Corollary on the behaviour of the martingales $\mathcal{M}_t^i$.

Corollary 3.2.8. *We have for* $\mathcal{M}_t^i = \int_0^t \sqrt{2\gamma^i X_s^i X_s^{i+1}} dB_s^i$ *that* $\left|\mathcal{M}_{\bar{t}}^i - \mathcal{M}_t^i\right| \to 0$ *for* $\bar{t} \geq t \to \infty$ *a.s.*

Proof. $\mathcal{M}_t^i$ is a continuous martingale with $d < \mathcal{M}^i >_t = 2\gamma^i X_t^i X_t^{i+1} dt$ and $\mathcal{M}_0^i = 0$. By [10], Theorem IV.(34.11) we can time-change $\mathcal{M}_t^i$ such that $\mathcal{M}_t^i = B_{<\mathcal{M}^i>_t}$, where B_t is a Brownian motion on a suitably extended probability space. Now

$$\begin{aligned} 0 \leq < \mathcal{M}^i >_{\bar{t}} - < \mathcal{M}^i >_t &= \int_t^{\bar{t}} 2\gamma^i X_s^i X_s^{i+1} ds \\ &\leq \int_t^{\bar{t}} \sum_{i=1}^d 2\gamma^i X_s^i X_s^{i+1} ds = < S >_{\bar{t}} - < S >_t \to 0 \text{ for } \bar{t} \geq t \to \infty \end{aligned}$$

as S_t converges a.s. and thus $\lim\limits_{t\to\infty} < S >_t < \infty$ a.s. Hence we obtain

$$\left|\mathcal{M}_{\bar{t}}^i - \mathcal{M}_t^i\right| = \left|B_{<\mathcal{M}^i>_{\bar{t}}} - B_{<\mathcal{M}^i>_t}\right| \to 0 \text{ for } \bar{t} \geq t \to \infty \text{ a.s.} \tag{3.8}$$

as required. □

Continuation of the proof of Lemma 3.2.7.

2. case. $\tau(\omega) < \eta(\omega) < \infty$.

Suppose $\tau < \eta < \infty$. This implies that there exists $\epsilon > 0$ such that $I|_{[\tau,\tau+\epsilon)} = 0$ and $S|_{[\tau,\tau+\epsilon)} > 0$. By definition of I_t, S_t and the continuity of the processes under consideration this requires that there exists $\delta = \delta(\omega) > 0$ and $i = i(\omega) \in \{1, \ldots, d\}$ such that

$$X_t^i|_{[\tau,\tau+\delta)} = 0 \text{ and } X_t^{i+1}|_{[\tau,\tau+\delta)} > \delta. \tag{3.9}$$

Next consider increments in the i^{th} coordinate to see that for all $0 \leq t < \delta$

$$\begin{aligned}
0 &= X_{\tau+t}^i - X_\tau^i \\
&= \left(\mathcal{M}_{\tau+t}^i - \mathcal{M}_\tau^i\right) + \int_\tau^{\tau+t} \sum_{j=1, j\neq i}^{d} \left(X_s^j q_{ji} - X_s^i q_{ij}\right) ds \\
&\geq \left(\mathcal{M}_{\tau+t}^i - \mathcal{M}_\tau^i\right) - (d-1) \max_{i\neq j} q_{ij} \int_\tau^{\tau+t} X_s^i ds + \int_\tau^{\tau+t} X_s^{i+1} q_{i+1,i} ds \\
&\geq \left(\mathcal{M}_{\tau+t}^i - \mathcal{M}_\tau^i\right) + t q_{i+1,i} \delta.
\end{aligned}$$

As $I|_{[\tau,\tau+\epsilon)} = 0$ and $d < S >_s = I_s ds$ we get $\left.\frac{d<S>_s}{ds}\right|_{[\tau,\tau+\epsilon)} = 0$. Similarly to the proof of Corollary 3.2.8 this implies $\left.\frac{d<\mathcal{M}^i>_s}{ds}\right|_{[\tau,\tau+\epsilon)} = 0$. Rewriting $\mathcal{M}_t^i = B_{<\mathcal{M}^i>_t}$ as done in the Corollary we get $\mathcal{M}^i|_{[\tau,\tau+\epsilon)} \equiv$ const. and thus $\mathcal{M}_{\tau+t}^i - \mathcal{M}_\tau^i = 0$. Plugging this in the above inequality we obtain $0 \geq t q_{i+1,i}\delta > 0$, a contradiction.

Taking both cases together we have shown that $\eta = \infty$ a.s., i.e. $\inf\{t \geq 0 : S_t = 0\} = \infty$ a.s. □

Remark 3.2.9. *We have actually shown more. As $\eta < \infty$ was not used in the proof of case 2 of Lemma 3.2.7, we have moreover shown that $\tau = \eta = \infty$.*

Also observe that the proof of Lemma 3.2.7 only uses $q_{ij} \geq 0, j \neq i$ and $q_{i+1,i} > 0$ for all $i \in \{1, \ldots, d\}$ and so in particular holds for nearest neighbour random walk on the circle, even though uniqueness for this case remains open.

Lemma 3.2.10. *The overall sum S_t of our processes converges to* 0 *a.s., i.e. $S_t \to 0$ for $t \to \infty$. As the processes $X_t^i \leq S_t$, $i \in \{1, \ldots, d\}$ are nonnegative, they converge to* 0 *a.s. as well.*

Proof. We shall prove the assertion by constructing a contradiction. The a.s.-existence of a limit S_∞ was shown in Corollary 3.2.6. Suppose by contradiction that $S_\infty = S_\infty(\omega) > 0$ for ω element of a set of positive measure. Choose $0 < \epsilon < S_\infty$ arbitrarily small. By (3.8) there exists $T = T(\epsilon, \omega) \geq 0$ such that for all $\bar{t} > t \geq T$

$$\left|\mathcal{M}_{\bar{t}}^i - \mathcal{M}_t^i\right| \leq \epsilon, i \in \{1, \ldots, d\}, \text{ and } |S_t - S_\infty| \leq \epsilon. \tag{3.10}$$

Now observe that for $\bar{t} > t$

$$\begin{aligned}
\left|X^i_{\bar{t}} - X^i_t\right| &\geq -\left|\mathcal{M}^i_{\bar{t}} - \mathcal{M}^i_t\right| + \int_t^{\bar{t}} \sum_{j=1, j\neq i}^{d} \left(X^j_s q_{ji} - X^i_s q_{ij}\right) ds \\
&\geq -\left|\mathcal{M}^i_{\bar{t}} - \mathcal{M}^i_t\right| + q_{min} \int_t^{\bar{t}} \sum_{j=1, j\neq i}^{d} X^j_s ds - (d-1) q_{max} \int_t^{\bar{t}} X^i_s ds \\
&\geq -\left|\mathcal{M}^i_{\bar{t}} - \mathcal{M}^i_t\right| + q_{min} \int_t^{\bar{t}} (S_s - X^i_s) ds - (d-1) q_{max} \int_t^{\bar{t}} X^i_s ds \\
&= -\left|\mathcal{M}^i_{\bar{t}} - \mathcal{M}^i_t\right| + q_{min} \int_t^{\bar{t}} S_s ds - \left[(d-1) q_{max} + q_{min}\right] \int_t^{\bar{t}} X^i_s ds.
\end{aligned}$$

Hence we have for $\bar{t} > t \geq T$

$$\left|X^i_{\bar{t}} - X^i_t\right| \geq -\epsilon + q_{min}(\bar{t} - t)\,(S_\infty - \epsilon) - \left[(d-1) q_{max} + q_{min}\right](\bar{t} - t) \sup_{s\in[t,\bar{t}]} X^i_s.$$

This is equivalent to

$$\sup_{s\in[t,\bar{t}]} X^i_s \geq \frac{-\epsilon + q_{min}(\bar{t} - t)\,(S_\infty - \epsilon) - \left|X^i_{\bar{t}} - X^i_t\right|}{\left[(d-1) q_{max} + q_{min}\right](\bar{t} - t)}. \tag{3.11}$$

In what follows fix $i \in \{1, \ldots, d\}$ arbitrary and use the following notation

$$0 \leq I \equiv \liminf_{t\to\infty} X^i_t \leq \limsup_{t\to\infty} X^i_t \equiv S \leq S_\infty < \infty.$$

We shall prove that $S_\infty > 0$ implies $I > 0$, which will provide us with the desired contradiction in the end.

1. case: $S_\infty > 0$ and $I < S$.
As X^i_t is a continuous process, there exists an increasing sequence $\{t_n\}_{n\in\mathbb{N}}$, $t_n \to \infty$ for $n \to \infty$, independent of the choice of ϵ, such that

$$X^i_{t_n} = I + \delta(S - I), n \in \mathbb{N},$$

where $0 < \delta < 1$ fixed but arbitrarily small (see Figure 3.1). Without loss of generality let

$$\sup_{s\in[t_{2n-1},t_{2n}]} X^i_s = X^i_{t_{2n-1}} = X^i_{t_{2n}} = I + \delta(S - I), n \in \mathbb{N}.$$

Applying (3.11) with $t = t_{2n-1}, \bar{t} = t_{2n}$ (choose $n \in \mathbb{N}$ such that $t \geq T$) gives

$$I + \delta(S - I) \geq \frac{-\epsilon + q_{min}(t_{2n} - t_{2n-1})\,(S_\infty - \epsilon)}{\left[(d-1) q_{max} + q_{min}\right](t_{2n} - t_{2n-1})}. \tag{3.12}$$

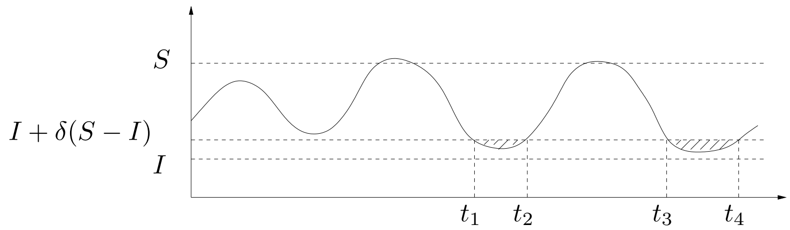

Figure 3.1: The definition of t_{2n-1} and t_{2n}.

As the choice of n may depend on ϵ we have to find an estimate for the term $\left|\frac{\epsilon}{t_{2n}-t_{2n-1}}\right|$ before considering $\epsilon \to 0^+$. First observe that for $0 \leq u < v$ to be specified later

$$\begin{aligned}\left|X_v^i - X_u^i\right| &\leq \left|\mathcal{M}_v^i - \mathcal{M}_u^i\right| + \int_u^v \sum_{j=1}^d \left|X_s^j q_{ji}\right| ds \\ &\leq \left|\mathcal{M}_v^i - \mathcal{M}_u^i\right| + (v-u) q_{max} \sup_{0 \leq s < \infty} S_s\end{aligned}$$

and thus as S_t converges a.s.

$$v - u \geq \frac{\left|X_v^i - X_u^i\right| - \left|\mathcal{M}_v^i - \mathcal{M}_u^i\right|}{q_{max} \sup\limits_{0 \leq s < \infty} S_s}.$$

Now (3.8) yields that for all $\theta > 0$ small there exists $T' = T'(\theta, \omega)$ (note that T' is independent of the choice of ϵ) such that for all $v > u \geq T'$

$$v - u \geq \frac{\left|X_v^i - X_u^i\right| - \theta}{q_{max} \sup\limits_{0 \leq s < \infty} S_s}. \tag{3.13}$$

Moreover, choose T' such that for all $t_{2n-1} \geq T'$ we have

$$\inf_{s \in [t_{2n-1}, t_{2n}]} X_s^i < I + \frac{\delta}{2}(S - I),$$

which is possible by definition of I, S and δ. We get

$$\exists u, v \in [t_{2n-1}, t_{2n}] \text{ s.t. } |X_v^i - X_u^i| > \frac{\delta}{2}(S - I) \overset{(3.13)}{\Rightarrow} v - u \geq \frac{\frac{\delta}{2}(S-I) - \theta}{q_{max} \sup\limits_{0 \leq s < \infty} S_s}.$$

By choosing θ sufficiently small we obtain that there exists $T' = T'(\theta, \omega, \delta)$ such that $|t_{2n} - t_{2n-1}| \geq v - u >$ const$(\omega, \delta) > 0$ for all $t_{2n-1} \geq T'$.

Let us return to (3.12). Letting $\epsilon \to 0^+$ yields

$$I + \delta(S - I) \geq \frac{q_{min} S_\infty}{[(d-1) q_{max} + q_{min}]}.$$

Now let $\delta \to 0^+$ to finally obtain $I > 0$.

2. case: $S_\infty > 0$ and $I = S$.
As $I = S$ is equivalent to X_t^i being convergent, we can choose T as defined in (3.10) to additionally satisfy

$$|X_s^i - I| \leq \epsilon, \ \forall s \geq T.$$

Now (3.11) gives with $\bar{t} > t = T$, $\bar{t}$ arbitrary

$$\begin{aligned} I + \epsilon \geq \sup_{s \in [T,\bar{t}]} X_s^i &\geq \frac{-\epsilon + q_{min}(\bar{t} - T)\,(S_\infty - \epsilon) - \left|X_{\bar{t}}^i - X_T^i\right|}{\left[(d-1)q_{max} + q_{min}\right](\bar{t} - T)} \\ &\geq \frac{-\epsilon + q_{min}(\bar{t} - T)\,(S_\infty - \epsilon) - 2\epsilon}{\left[(d-1)q_{max} + q_{min}\right](\bar{t} - T)}. \end{aligned}$$

Taking $\epsilon \to 0^+$ yields $I > 0$ once more.

Taking both cases together we have shown $\liminf_{t\to\infty} X_t^i > 0$ for all $i \in \{1, \ldots, d\}$, given $S_\infty > 0$, as i was chosen arbitrary in the calculations. As the X_t^i are continuous processes this gives a contradiction to
$\lim_{t\to\infty} < S >_t = \int_0^\infty \sum_{i=1}^d 2\gamma^i X_s^i X_s^{i+1} ds < \infty$ a.s., as this requires that the limit inferior of the nonnegative integrand becomes zero for $t \to \infty$. $S_\infty > 0$ thus gives a contradiction, i.e. we have shown $S_\infty = 0$ a.s. □

Lemma 3.2.11. *The solution to the given SDE (3.2) is unique in law for all* $X_0 \in \mathbb{R}^d$ *s.t.* $X_0^i \geq 0, i \in \{1, \ldots, d\}$ *and* $\prod_{i=1}^d \left(X_0^i + X_0^{i+1}\right) > 0$.

Proof. We shall apply Theorem 4 in [3] (or alternatively Theorem 1.6 of [9]). The Hypotheses under which this Theorem is stated hold, except for condition (3) in Hypothesis 2. Here we have that $b_i(x) > 0$ if $x_i = 0$, except for the case where $x = 0 \in \mathbb{R}^d$ (here $b_i(0) = 0$). Thus we have to consider this case separately.

By modifying the drift coefficients in a small open neighbourhood of 0, say in $B_\epsilon(0)$ with $\epsilon > 0$ arbitrarily small, we can achieve that the drift coefficients satisfy all conditions such as Hölder continuity on compact subsets of $\mathbb{R}_+^d$ and $|b_i(x)| \leq c(1 + |x|)$ and additionally are strictly positive on $B_\epsilon(0)$. For instance, let

$$\tilde{b}_i(x) \equiv b_i(x) + (\epsilon - |x|) \vee 0.$$

By solving the system with these modified coefficients we obtain existence and uniqueness of the modified solution. Finally take $\epsilon > 0$ so small that the starting point $x_0 \notin B_\epsilon(0)$. As the diffusion and drift coefficients of the modified SDE are identical with the ones of the original SDE on $(B_\epsilon(0))^c$, every solution to (3.2) is unique until it hits $B_\epsilon(0)$. By taking $\epsilon \downarrow 0^+$ and recalling that we showed that every solution to (3.2) does not hit 0 in finite time (see Lemma 3.2.7) we obtain the assertion. □

3.2.3 The normalized processes

Corollary 3.2.12. *The corresponding SDEs for the normalized processes*

$$Y_t^i \equiv \frac{X_t^i}{S_t} \tag{3.14}$$

with $0 \le Y_t^i \le 1$ *are as follows.*

$$\begin{aligned} dY_t^i = &- Y_t^i \sum_{j \neq i} \sqrt{2\gamma^j Y_t^j Y_t^{j+1}} dB_t^j + (1 - Y_t^i) \sqrt{2\gamma^i Y_t^i Y_t^{i+1}} dB_t^i \\ &+ Y_t^i \sum_{j \neq i} 2\gamma^j Y_t^j Y_t^{j+1} dt + (Y_t^i - 1) 2\gamma^i Y_t^i Y_t^{i+1} dt + \sum_{j=1}^d Y_t^j q_{ji} dt. \end{aligned} \tag{3.15}$$

Idea of the proof. The proof is an easy application of Itô's formula. □

In what follows we shall consider the above system of SDEs for the Y_t^i without referring to their derivation via the X_t^i and ask for existence and uniqueness of solutions. As we have not shown nonnegativity of the Y_t^i yet, we replace the terms under the square root with their absolute values.

Proposition 3.2.13. *The SDE (3.15) started at* $Y_0 \in [0,1]^d \backslash \partial [0,1]^d$ *with* $\sum_i Y_0^i = 1$ *has a unique in law solution. Moreover the solution satisfies* $Y_t \in [0,1]^d$ *and* $\sum_i Y_t^i = 1$ *for all* $t \ge 0$ *a.s.*

Proof. Existence follows immediately from the existence of solutions X_t to (3.2). Indeed, let $X_0 \equiv Y_0$, then $Y_t^i \equiv X_t^i / \sum_j X_t^j$ solves (3.15) by Corollary 3.2.12 with $Y_0^i = X_0^i / \sum_j X_0^j = Y_0^i / \sum_j Y_0^j = Y_0^i$.

As in Corollary 3.2.2 one can show that the local times at zero of the processes Y_t^i are zero. The nonnegativity of the processes Y_t^i can be shown as follows.

Lemma 3.2.14. *The processes* Y_t^i, $i \in \{1, \ldots, d\}$ *are nonnegative.*

Proof. We have for all $1 \le i \le d$ and $t \le T_M = T_M(\omega) \equiv \inf\{t \ge 0 : \max_i |Y_t^i| \ge M\}$, $M \in \mathbb{N}$ fixed,

$$\begin{aligned} \left(Y_t^i\right)^- &= \int_0^t -1_{\{Y_s^i \le 0\}} dY_s^i \\ &= \text{mart.} + \int_0^t -1_{\{Y_s^i \le 0\}} \Bigg(Y_s^i \sum_{j \neq i} 2\gamma^j Y_s^j Y_s^{j+1} + (Y_s^i - 1) 2\gamma^i Y_s^i Y_s^{i+1} \\ &\qquad + \sum_{j=1}^d Y_s^j q_{ji} \Bigg) ds. \end{aligned}$$

We can bound the above by

$$\begin{aligned}
&\text{mart.} + \int_0^t C(M)\left(Y_s^i\right)^- - 1_{\{Y_s^i \le 0\}} \sum_{j\neq i} Y_s^j q_{ji} ds \\
&\le \text{mart.} + \int_0^t C(M)\left(Y_s^i\right)^- + C\sum_{j\neq i}\left(Y_s^j\right)^- ds.
\end{aligned}$$

Taking expectations on both sides and summing over all coordinates we get for all $t \ge 0$,

$$\begin{aligned}
\sum_i \mathbb{E}\Big[\left(Y_{t\wedge T_M}^i\right)^-\Big] &\le \int_0^t C(M)\sum_i \mathbb{E}\Big[\left(Y_{s\wedge T_M}^i\right)^-\Big] + C\sum_i\sum_{j\neq i}\mathbb{E}\Big[\Big(Y_{s\wedge T_M}^j\Big)^-\Big] ds \\
&\le C(M)\int_0^t \sum_i \mathbb{E}\Big[\left(Y_{s\wedge T_M}^i\right)^-\Big] ds.
\end{aligned}$$

An application of Gronwall's lemma yields $\mathbb{E}\Big[\left(Y_{t\wedge T_M}^i\right)^-\Big] = 0$. Take $M \to \infty$ and use Fatou's lemma to obtain the claim. $\square$

As we shall show in the following Corollary $\sum_i Y_0^i = 1$ implies $\sum_i Y_t^i = 1$ for all $t \ge 0$.

Corollary 3.2.15. *Every solution $(Y_t^i)_{1\le i\le d}$ to (3.15) with $\sum_i Y_0^i = 1$ has to satisfy $\sum_i Y_t^i = 1$ for all $t \ge 0$ a.s.*

Proof. We have from (3.15) and (3.3) that

$$\begin{aligned}
d\left(\sum_{i=1}^d Y_t^i\right) &= \left(1 - \sum_{i=1}^d Y_t^i\right)\sum_{j=1}^d\left[\sqrt{2\gamma^j Y_t^j Y_t^{j+1}}dB_t^j - 2\gamma^j Y_t^j Y_t^{j+1}dt\right] \\
&= \left(1 - \sum_{i=1}^d Y_t^i\right)\left(d\mathcal{N}_t - d < \mathcal{N} >_t\right),
\end{aligned}$$

where $\mathcal{N}_t \equiv \int_0^t \sum_{j=1}^d \sqrt{2\gamma^j Y_s^j Y_s^{j+1}}dB_s^j$ and $\mathcal{N}_{t\wedge T_M}$ is a martingale starting at 0.

Setting $D_t \equiv \sum_{i=1}^d Y_t^i$ and applying Itô's formula we obtain $(D_0 - 1 = 0)$

$$\begin{aligned}
&\left(D_{t\wedge T_M} - 1\right)^2 \\
&= \text{ mart. } + \int_0^{t\wedge T_M}\left\{2(D_s - 1)(-1 + D_s) + \frac{1}{2}2(1 - D_s)^2\right\} d < \mathcal{N} >_s
\end{aligned}$$

giving

$$\mathbb{E}(D_{t\wedge T_M} - 1)^2 \le 3C(M)\int_0^t \mathbb{E}(D_{s\wedge T_M} - 1)^2 ds.$$

Here we used that $d < \mathcal{N} >_s \le C(M)ds$ for $s \le t\wedge T_M$. Now Gronwall's lemma yields $D_t - 1 \equiv 0$ for all $t \le T_M$ a.s. Take $M \to \infty$ to obtain the claim. $\square$

Continuation of the proof of Proposition 3.2.13.
It remains to prove the uniqueness of solutions to (3.15). Observe that Corollary 3.2.15 implies $0 \leq Y_t^i \leq 1$ by the nonnegativity of the Y_t^i.

Now suppose that Y_t is a solution to (3.15) with $Y_0 \in [0,1]^d \backslash \partial[0,1]^d$ and such that $\sum_i Y_0^i = 1$. The following Lemma gives the existence of processes X_t^i such that $Y_t^i = X_t^i / \sum_i X_t^i$. This will enable us later to derive uniqueness of solutions to (3.15) from uniqueness of solutions to (3.2).

Lemma 3.2.16. *Given a process $(Y_t^i)_{1\leq i\leq d}$ that satisfies (3.15) with $Y_0 \in [0,1]^d \backslash \partial[0,1]^d$ and with $\sum_i Y_0^i = 1$, we can find a system of processes $(X_t^i)_{1\leq i\leq d}$ that satisfies (3.2) and (3.14) with $S_t \equiv \sum_i X_t^i$.*

Proof. We start with a motivation for the proof. Let $(Y_t^i)_{1\leq i\leq d}$ be given by (3.14). Definition (3.14) and $S_t = \sum_i X_t^i$ implied in the former setting that

$$dS_t = \sum_{i=1}^d dX_t^i = \sum_{i=1}^d \sqrt{2\gamma^i X_t^i X_t^{i+1}} dB_t^i = S_t \sum_{i=1}^d \sqrt{2\gamma^i Y_t^i Y_t^{i+1}} dB_t^i,$$

i.e.

$$dS_t \equiv S_t d\mathcal{M}_t \tag{3.16}$$

being solved by

$$S_t = S_0 \exp\left\{\mathcal{M}_t - \frac{1}{2} < \mathcal{M} >_t\right\}.$$

In the given setting the above calculation may be taken as a motivation to define

$$R_t \equiv R_0 \exp\left\{\mathcal{M}_t - \frac{1}{2} < \mathcal{M} >_t\right\}, \text{ where } \mathcal{M}_t \equiv \int_0^t \sum_{i=1}^d \sqrt{2\gamma^i Y_s^i Y_s^{i+1}} dB_s^i,$$

such that (3.16) holds with S_t replaced by R_t and $R_0 > 0$ to be chosen arbitrarily. Let $T_M \equiv \inf\{t \geq 0 : \max_i Y_t^i \geq M\}$. For $t \leq T_M$ we have $< \mathcal{M} >_t < \infty$ a.s. This finally leads to the definition of $(X_t^i)_{1\leq i\leq d}$, for $t \leq T_M$, by

$$X_t^i \equiv Y_t^i R_t = Y_t^i R_0 \exp\left\{\mathcal{M}_t - \frac{1}{2} < \mathcal{M} >_t\right\}, \tag{3.17}$$

where $X_0^i \equiv Y_0^i R_0$.

Let us check that $(X_t^i)_{1\leq i\leq d}$ as defined in (3.17) satisfies (3.2). Indeed, we have

$$\begin{aligned} dX_t^i &\overset{(3.17)}{=} R_t dY_t^i + Y_t^i dR_t + < Y^i, R >_t \\ &\overset{(3.15),(3.16)}{=} R_t \sqrt{2\gamma^i Y_t^i Y_t^{i+1}} dB_t^i + R_t \sum_{j=1}^d Y_t^j q_{ji} dt \\ &\overset{(3.17)}{=} \sqrt{2\gamma^i X_t^i X_t^{i+1}} dB_t^i + \sum_{j=1}^d X_t^j q_{ji} dt, \end{aligned}$$

which proves our claim after taking $M \to \infty$ and observing that by Corollary 3.2.15, $R_t = \sum_i Y_t^i R_t = \sum_i X_t^i$. □

Conclusion of the proof of Proposition 3.2.13.
The uniqueness of the solution to (3.15) follows from the uniqueness of the solution $(X_t^i)_{1\le i\le d}$ to (3.2). Indeed, observe that the uniqueness of X_t yields the uniqueness of $R_t = \sum_i Y_t^i R_t = \sum_i X_t^i$ and thus of $Y_t^i = X_t^i/R_t$. □

3.2.4 Properties of a stationary distribution to the system (3.15) of normalized processes

Recall that we are given the system of SDEs (3.15) and look for solutions satisfying $\sum_i Y_t^i = 1$, where $0 \le Y_t^i \le 1$, $1 \le i \le d$.

In what follows we shall look for stationary distributions to this system. By Proposition IV.9.2 of [7], a measure π is a stationary distribution for our process, that is, it is a stationary distribution for $\mathcal{A}$, where $\mathcal{A}$ denotes the generator of our system of SDEs, if and only if $\int \mathcal{A}g d\pi = 0$ for all $g \in \mathcal{C}^2$. Hence we shall investigate necessary properties of a measure π satisfying $\int_{[0,1]^d} \mathcal{A}g d\pi = 0$ for all $g \in \mathcal{C}^2$. In particular we want to show the following Proposition.

Proposition 3.2.17. *If π is stationary then it does not put mass on the set*

$$\mathcal{N} \equiv \{y \in [0,1]^d : \exists i \in \{1,\dots,d\} : y_i = 0\},$$

i.e. on the set where at least one of the coordinate processes becomes extinct.

Proof. The generator $\mathcal{A}$ of our diffusion Y can be determined to be

$$\mathcal{A}g(x) = \sum_{i=1}^d \partial_i g(x) b_i(x) + \frac{1}{2}\sum_{i,j=1}^d \partial_{ij} g(x) a_{ij}(x), \tag{3.18}$$

where

$$b_i(x) = x_i \sum_{j\ne i} 2\gamma^j x_j x_{j+1} + (x_i - 1)2\gamma^i x_i x_{i+1} + \sum_{j=1}^d x_j q_{ji}, \tag{3.19}$$

$$a_{ii}(x) = (x_i)^2 \sum_{j\ne i} 2\gamma^j x_j x_{j+1} + (1-x_i)^2\, 2\gamma^i x_i x_{i+1} \text{ and for } i \ne j$$

$$a_{ij}(x) = x_i x_j \sum_{k\notin\{i,j\}} 2\gamma^k x_k x_{k+1} - (1-x_i)x_j 2\gamma^i x_i x_{i+1} - x_i(1-x_j)2\gamma^j x_j x_{j+1}$$

(see [10], V.(1.7) for the definition of $\mathcal{A}$).

In what follows we shall try to find a function $g \in C^2$ which leads to a contradiction to $\int \mathcal{A}g d\pi = 0$ in case π puts mass on the set $\mathcal{N}$. Thereby we shall take advantage of the observation that for $x_i = 0$ for $i \in \{1,\dots,d\}$ arbitrarily fixed we have

$$a_{ii}(x) = 0 \text{ and } b_i(x) \ge \sum_{j\ne i} x_j q_{min} = (1-x_i) q_{min} = q_{min} > 0, \tag{3.20}$$

where we set $q_{min} \equiv \min_{i \neq j} q_{ij}$.

To make things easier we shall fix $i \in \{1, \dots, d\}$ arbitrarily and only look for functions $g(x) = g(x_i), g \in \mathcal{C}^2$. Thus we obtain for (3.18) with $f(x_i) = f(x) \equiv \partial_i g(x) = \partial_i g(x_i), f \in \mathcal{C}^1$ that $\int_{[0,1]^d} \mathcal{A} g d\pi = 0$ is equivalent to

$$\int_{[0,1]^d} f(x_i) \left[x_i \sum_{j \neq i} 2\gamma^j x_j x_{j+1} + (x_i - 1) 2\gamma^i x_i x_{i+1} + \sum_{j=1}^{d} x_j q_{ji} \right] d\pi(x)$$
$$= - \int_{[0,1]^d} \frac{1}{2} \partial_i f(x_i) \left[(x_i)^2 \sum_{j \neq i} 2\gamma^j x_j x_{j+1} + (1 - x_i)^2 \, 2\gamma^i x_i x_{i+1} \right] d\pi(x).$$

Using that $x \in [0,1]^d$ we get

$$\int_{[0,1]^d} \left\{ - |f(x_i)| \, C_1 x_i + f(x_i) \sum_{j \neq i} x_j q_{ji} \right\} d\pi(x) \leq C_2 \int_{[0,1]^d} |\partial_i f(x_i)| \, x_i d\pi(x),$$

where all constants under consideration are nonnegative. Assuming that f is nonnegative and as $\sum_{j \neq i} x_j q_{ji} \geq (1 - x_i) q_{min}$ (see (3.20)) we finally get

$$\int_{[0,1]^d} f(x_i) \left[-C_1 x_i + (1 - x_i) q_{min} \right] d\pi(x) \leq C_2 \int_{[0,1]^d} |\partial_i f(x_i)| \, x_i d\pi(x). \quad (3.21)$$

In what follows we shall try to find a nonnegative function $f \in \mathcal{C}^1$ which gives a contradiction to the assumption that π puts mass on the set $\mathcal{N}$.

As we want to investigate the behaviour of π on the set $\mathcal{N}$, we shall define a function $f \in \mathcal{C}^1$ with support in $[0,1]$ and then "squeeze" the function, i.e. rescale it in such a way that the support of the new function f_ϵ lies in $[0, \epsilon]$. This way we localize equation (3.21) at $x_i = 0$. Let us make this more precise.

Suppose we are given

$$f \in \mathcal{C}^1_+ \text{ with support in } [0,1] \text{ or } (-\infty, 1]$$

and let $f_\epsilon(x) \equiv f\left(\frac{x}{\epsilon}\right)$ then

$$f_\epsilon \in \mathcal{C}^1_+ \text{ with support in } [0,\epsilon] \text{ or } (-\infty, \epsilon] \text{ and } f'_\epsilon(x) = f'\left(\frac{x}{\epsilon}\right) \frac{1}{\epsilon}.$$

Choose for instance $f(x) = \exp\left(-\frac{1}{1-x}\right) 1(x \leq 1)$. Plugging this into (3.21) and abbreviating $A_\epsilon \equiv [0, \epsilon] \times [0,1]^{d-1}$ we obtain for arbitrary $0 < \epsilon < 1$

$$\int_{A_\epsilon} f\left(\frac{x}{\epsilon}\right) \left[-C_1 \epsilon + (1 - \epsilon) q_{min} \right] d\pi(x, z) \leq \int_{A_\epsilon} \left| f'\left(\frac{x}{\epsilon}\right) \right| C_2 \frac{x}{\epsilon} d\pi(x, z),$$

where we assumed without loss of generality that $i = 1$. This yields

$$\int_{A_\epsilon} \exp\biggl(-\frac{1}{1-\frac{x}{\epsilon}}\biggr) \left\{ [-C_1\epsilon + (1-\epsilon)q_{min}] - \frac{1}{\left(1-\frac{x}{\epsilon}\right)^2} \frac{x}{\epsilon} C_2 \right\} d\pi(x,z) \quad (3.22)$$
$$\equiv I_1(\epsilon) + I_2(\epsilon) \leq 0$$

for all $\epsilon > 0$.

For the first part I_1 of the integral observe that the absolute value of the integrand is bounded for $0 < \epsilon \leq \epsilon_0$ via $C_1\epsilon_0 + q_{min}$. Hence we can apply the dominated convergence theorem to the first integral to obtain

$$\lim_{\epsilon \downarrow 0^+} I_1(\epsilon) = e^{-1} q_{min} \pi\bigl(\{0\} \times [0,1]^{d-1}\bigr) . \quad (3.23)$$

For the second part of the integral I_2 note that for $x \in [0,\epsilon]$ the absolute value of the integrand is bounded by $4e^{-2}\frac{x}{\epsilon}C_2 \leq 4e^{-2}C_2$. As this result is uniform in ϵ we can apply the dominated convergence theorem again to obtain $\lim\limits_{\epsilon \downarrow 0^+} I_2(\epsilon) = 0$.

Plugging this and (3.23) into (3.22) we get with $q_{min} > 0$

$$e^{-1} q_{min} \pi\bigl(\{0\} \times [0,1]^{d-1}\bigr) \leq 0 \Rightarrow \pi\bigl(\{0\} \times [0,1]^{d-1}\bigr) = 0,$$

i.e. π does not put mass on $\mathcal{N}$ as stated above. □

3.2.5 Stationary distribution

Recall that we are given the system of SDEs (3.15) which we rewrite as

$$dY_t = \sigma(Y_t)dB_t + b(Y_t)dt$$

and look for solutions satisfying $\sum_i Y_t^i = 1$, where $0 \leq Y_t^i \leq 1$, $1 \leq i \leq d$.

Proposition 3.2.18. *The above system of SDEs has a unique stationary distribution π supported by $S \equiv \bigl([0,1]^d \backslash \partial[0,1]^d\bigr) \cap \{y : \sum_i y_i = 1\}$. Moreover, $\mathcal{L}(Y_t|Y_0 = y) \Rightarrow \pi$ holds for all $y \in S$.*

Proof. Let Y_t be the unique strong Markov solution to (3.15). Recall that we already showed that every equilibrium distribution for $\mathcal{A}$ doesn't put mass on $\mathcal{N} = \{y : \exists i : y_i = 0\}$ in Proposition 3.2.17 and that $\sum_i Y_t^i = 1$ for all $t \geq 0$ in Proposition 3.2.13. Hence, if Y_t has any equilibrium distributions, they are concentrated on

$$S \equiv \bigl([0,1]^d \backslash \partial[0,1]^d\bigr) \cap \left\{ y : \sum_i y_i = 1 \right\} .$$

In what follows we shall consider the process $\tilde{Y}_t \equiv (Y_t^1, \ldots, Y_t^{d-1}) \in [0,1]^{d-1}$ instead. The martingale problem for the resulting SDE for $\tilde{Y}$ is consequently well-posed as the corresponding martingale problem for Y is well-posed. For

$$\tilde{x} \in \tilde{S} \equiv \left([0,1]^{d-1} \cap \left\{\sum_{i=1}^{d-1} \tilde{x}_i \leq 1\right\}\right) \setminus \left\{\tilde{x} : \exists i : \tilde{x}_i = 0 \text{ or } \sum_{i=1}^{d-1} \tilde{x}_i = 1\right\}$$
$$= \left([0,1]^{d-1} \setminus \partial [0,1]^{d-1}\right) \cap \left\{\tilde{x} : 0 < \sum_i \tilde{x}_i < 1\right\},$$

$\tilde{a}_{ij}(\tilde{x})$ is non-singular by Corollary A.1.1 of the Appendix. Also observe that $\tilde{S}$ is an open subset of $[0,1]^{d-1}$ compact. Now the reasoning of [2], Section 3.1 can be applied to show that the system of SDEs for $\tilde{Y}$ has a unique stationary distribution $\tilde{\pi}$ supported by $\tilde{S}$ and that $\mathcal{L}(\tilde{Y}_t|\tilde{Y}_0 = \tilde{y}) \Rightarrow \tilde{\pi}$ holds for all $\tilde{y} \in \tilde{S}$. Note that, as in [2], the non-singularity of $\tilde{a}_{ij}(\tilde{x})$ on $\tilde{S}$ is crucial.

The claim now follows from $\tilde{Y}_t \equiv (Y_t^1, \ldots, Y_t^{d-1})$ and $Y_t^d = 1 - \sum_{i=1}^{d-1} Y_t^i$.

A complete proof is given in the Appendix, Subsection A.2. □

3.2.6 Extension to arbitrary networks

Instead of (3.1) we can consider

$$dX_t^i = \sqrt{2\gamma^i X_t^i \sum_{j \in C_i} X_t^j} dB_t^i + \sum_{j=1}^d X_t^j q_{ji} dt, \; i \in \{1, \ldots, d\}, \tag{3.24}$$

where $C_i \subset \{1, \ldots, d\}$ with $i \notin C_i$ and $|C_i| \geq 1$. We can think of C_i as the set of catalysts of i. The cyclic case corresponds to $C_i = \{i+1\}$. We shall assume as above that γ^i and $q_{ji}, i \neq j$ are given positive constants and the $X_0^i \geq 0, i \in \{1, \ldots, d\}$ are given initial conditions. (q_{ji}) is again a Q-matrix modelling mutations from type j to type i.

For this setup the above proofs directly carry over. Observe in particular that (3.9) changes to the requirement that there exists $\delta = \delta(\omega) > 0$ and $i = i(\omega), j = j(\omega) \in \{1, \ldots, d\}$ such that

$$X_t^i|_{[\tau, \tau+\delta)} = 0 \text{ and } X_t^j|_{[\tau, \tau+\delta)} > \delta.$$

Also note that the restriction on the state space for the initial condition in Lemma 3.2.11 changes to $\prod_{i=1}^d \left(X_0^i + \sum_{j \in C_i} X_0^j\right) > 0$ as we now use Theorem 1.6 of [9].

3.2.7 Complete analysis of the case $d = 2$

Remark 3.2.19. *We shall denote γ^i by γ_i in what follows, as for instance γ^2 might easily be misunderstood.*

Recall that the normalized processes $Y_t^i = \frac{X_t^i}{S_t}$ for our given SDE satisfy (3.15).

Corollary 3.2.20. *For $d = 2$ we obtain the following SDE for $Y_t \equiv Y_t^1$ (note that $Y_t^2 = 1 - Y_t^1$)*

$$
\begin{aligned}
dY_t =& \sqrt{2(1-Y_t)Y_t\left((Y_t)^2\gamma_2 + (Y_t-1)^2\gamma_1\right)}dB_t \qquad (3.25)\\
&+ 2(1-Y_t)Y_t\left(Y_t\gamma_2 + (Y_t-1)\gamma_1\right)dt + Y_t\left(q_{11}-q_{21}\right)dt + q_{21}dt.
\end{aligned}
$$

Idea of the proof. We can calculate the SDE for Y_t^1 from (3.15), using our cyclic definition $Y_t^{2+1} = Y_t^1$. We get in particular that

$$d < Y^1 >_t = 2(1-Y_t^1)Y_t^1\left((Y_t^1)^2\gamma_2 + (Y_t^1-1)^2\gamma_1\right)dt.$$

Hence we can rewrite Y_t^1 on a possibly enlarged probability space in terms of a Brownian motion B_t as above (cf. [10], Theorem IV.(34.11)). □

In what follows we shall prove existence and pathwise uniqueness of solutions to the SDE (3.25) under the constraint $Y_t \in [0,1]$. First observe that (3.25) implies that

$$dY_t = \sigma(Y_t)dB_t + b(Y_t)dt, \qquad (3.26)$$

where $Y_t \in I \equiv [0,1]$ with

$$\sigma(x) = \sqrt{2(1-x)x\left(x^2\gamma_2 + (x-1)^2\gamma_1\right)} \qquad (3.27)$$

and

$$b(x) = 2(1-x)x\left(x\gamma_2 + (x-1)\gamma_1\right) + x\left(q_{11}-q_{21}\right) + q_{21}.$$

Lemma 3.2.21. *If $Y_0 \in [0,1]$, the SDE (3.26) has a pathwise unique solution, taking values in $[0,1]$.*

Proof. Replace $\sigma(x)$ and $b(x)$ in (3.26) by continuous functions $\tilde{\sigma}(x), \tilde{b}(x)$ with compact support such that they coincide with $\sigma(x), b(x)$ on $[0,1]$. Then existence of solutions to the SDE with modified coefficients follows by Theorem V.3.10 in [7].

By reasoning as in Lemma 3.2.14 (first consider $(Y_t)^-$, then $(1-Y_t)^-$) we can further show that any solution to the modified SDE satisfies $0 \le Y_t \le 1$ a.s. and is therefore a solution to the given SDE as well.

Pathwise uniqueness of a solution follows from the Yamada-Watanabe pathwise-uniqueness theorem for 1-dim. diffusions (see V.(40.1) in [10] and replace the term under the square root in (3.27) by its absolute value). □

As every one-dimensional diffusion can be uniquely characterized by its scale function and speed measure, we shall calculate the scale function as a first step towards investigating the long-time behaviour of the given one-dimensional SDE for $Y_t = Y_t^1$.

Lemma 3.2.22. *The scale function of Y is given by (up to increasing affine transformations)*

$$
\begin{aligned}
s'(x) =& \left|x^2\gamma_2 + (x-1)^2\gamma_1\right|^{-\left(1+\frac{q_{11}}{2\gamma_2}-\frac{q_{21}}{2\gamma_1}\right)} |1-x|^{\frac{q_{11}}{\gamma_2}} |x|^{-\frac{q_{21}}{\gamma_1}}\\
&\times \exp\left\{-\left(\frac{q_{11}+q_{21}}{\sqrt{\gamma_1\gamma_2}}\right)\arctan\left[\frac{(\gamma_1+\gamma_2)x-\gamma_1}{\sqrt{\gamma_1\gamma_2}}\right]\right\}.
\end{aligned}
$$

This yields in particular that $s'(0) = s'(1) = \infty$*. We obtain moreover*

$$s(0) = \begin{cases} const. > -\infty, & \frac{q_{21}}{\gamma_1} < 1, \\ -\infty, & o.w. \end{cases} \quad \text{and } s(1) = \begin{cases} const. < \infty, & \frac{-q_{11}}{\gamma_2} < 1, \\ \infty, & o.w. \end{cases}$$

Idea of the proof. Calculate the scale function as in [10], chapter V.28. □

Proposition 3.2.23. *We shall show the following result on speed and scale.*

(i) 0 *is* $\left\{\begin{array}{c} recurrent \\ never\ hit \\ recurrent \\ never\ hit \end{array}\right\}$ *and* 1 *is* $\left\{\begin{array}{c} recurrent \\ recurrent \\ never\ hit \\ never\ hit \end{array}\right\}$ *for* $\left\{\begin{array}{c} -q_{11} < \gamma_2, q_{21} < \gamma_1, \\ -q_{11} < \gamma_2, q_{21} \geq \gamma_1, \\ -q_{11} \geq \gamma_2, q_{21} < \gamma_1, \\ -q_{11} \geq \gamma_2, q_{21} \geq \gamma_1, \end{array}\right\}$,

where "0 is recurrent" should mean $P(\exists n \text{ such that } Y_t \neq 0 \text{ for all } t > n) = 0$ *and "1 never hit" that* $P(\exists t \geq 0 : Y_t = 1) = 0$*.*

(ii) In all cases $\int_0^\infty 1_{\{0,1\}}(Y_t)dt = 0$*.*

(iii) The scale function of Y *is given (up to increasing affine transformations) as in Lemma 3.2.22.*

(iv) The speed measure of the diffusion $Z \equiv s(Y)$ *in natural scale on* $[s(0), s(1)]$ *(read this as* $[s(0), s(1)] = (-\infty, s(1)]$ *for* $-\infty = s(0) < s(1) < \infty$ *etc.) is*

$$\begin{aligned} m(dz) &= e^{2\int_{\frac{1}{2}}^{s^{-1}(z)} 2b(u)\sigma(u)^{-2}du} \sigma(s^{-1}(z))^{-2} 1_{\{z \in (s(0), s(1))\}} dz \\ &= \frac{1}{(s'\sigma)^2} \circ s^{-1}(z) 1_{\{z \in (s(0), s(1))\}} dz. \end{aligned}$$

In particular, m *puts no mass on the endpoints* $s(0)$ *and* $s(1)$*.*

Idea of the proof. We shall mimic the calculations of [10], V.48, p. 287f. □

Corollary 3.2.24. *We obtain as a result on the limiting distribution the following. Let* $\{P_t\}$ *be the transition function of our diffusion in natural scale on* $\mathbb{R}$*. Then for each* x*:*

$$\lim_{t \to \infty} \| \pi - P_t(x, \cdot) \| = 0,$$

where

$$\pi(dz) \equiv \frac{m(dz)}{m(\mathbb{R})}.$$

Here $\|\cdot\|$ *denotes the total variation norm of a measure.*

Idea of the proof. This follows easily from [10], Theorem V.54.5. □

Bibliography

[1] Bass, R.F. and Perkins, E.A. Degenerate stochastic differential equations arising from catalytic branching networks. *Electron. J. Probab.* (2008) **13**, 1808–1885. MR2448130

[2] Dawson, D.A. and Greven, A. and den Hollander, F. and Sun, R. and Swart, J.M. The renormalization transformation for two-type branching models. *Ann. Inst. H. Poincaré Probab. Statist.* (2008) **44**, 1038–1077. MR2469334

[3] Dawson, D.A. and Perkins, E.A. On the uniqueness problem for catalytic branching networks and other singular diffusions. *Illinois J. Math.* (2006) **50**, 323–383 (electronic). MR2247832

[4] Eigen, M. and Schuster, P. *The Hypercycle: a principle of natural self-organization.* Springer, Berlin, 1979.

[5] Fleischmann, K. and Xiong, J. A cyclically catalytic super-brownian motion. *Ann. Probab.* (2001) **29**, 820–861. MR1849179

[6] Hofbauer, J. and Sigmund, K. *The Theory of Evolution and Dynamical Systems.* London Math. Soc. Stud. Texts, vol. 7, Cambridge Univ. Press, Cambridge, 1988. MR1071180

[7] Ethier, S.N. and Kurtz, T.G. *Markov Processes: Characterization and Convergence.* Wiley and Sons, Inc., Hoboken, New Jersey , 2005. MR0838085

[8] Karatzas, I. and Shreve, S.E. *Brownian Motion and Stochastic Calculus*, second edition. Springer, New York, 1991. MR1121940

[9] Kliem, S. Degenerate Stochastic Differential Equations for Catalytic Branching Networks. *To appear in Ann. Inst. H. Poincaré Probab. Statist.*

[10] Rogers, L.C.G. and Williams, D. *Diffusions, Markov Processes, and Martingales, vol. 2*, Reprint of the second (1994) edition. Cambridge Mathematical Univ. Press, Cambridge, 2000. MR1780932

Chapter 4

Convergence of Rescaled Competing Species Processes to a Class of SPDEs[1]

4.1 Introduction

We investigate convergence of certain rescaled models that have their applications in biology. Such convergence results can for instance be used to relate the limits to questions of coexistence and survival of types in the original models.

We start by introducing the underlying models and concepts for our later definitions in Subsections 4.1.1, 4.1.2 and 4.1.3. In Subsection 4.1.4 an overview of the results of this paper follows. Finally, in Subsection 4.1.5 we outline the remaining parts of the paper.

4.1.1 The voter model and the Lotka-Volterra model

An extensive introduction to the voter model can be found in Liggett [7], Chapter V. In short, the 1-dimensional voter model is a process $\xi_t : \mathbb{Z} \to \{0, 1\}$ with the following interpretation. $x \in \mathbb{Z}$ is seen as an individual with political opinion 0 or 1. This is the common interpretation which gives the model its name. Alternatively we can think of $\mathbb{Z}$ as space occupied by two populations 0 and 1. If $\xi_t(x) = 0$, at time t, the coordinate x is occupied by an individual of population 0. As we shall consider approximate densities later on, this interpretation will suit our purpose better in what follows.

The evolution of the process in time is given via infinitesimal rates. Following the notation in [7], let $c(x, \xi)$ denote the rate at which the coordinate $\xi(x)$ flips from 0 to 1 or from 1 to 0 when the system is in state ξ. Then the process ξ_t will satisfy

$$\mathbb{P}(\xi_t(x) \neq \xi_0(x)) = c(x, \xi_0)t + o(t) \text{ for } t \downarrow 0^+.$$

For the voter model, the rates can for instance be given by a random walk kernel

[1] A version of this chapter will be submitted for publication. Kliem, S.M. (2009) Convergence of Rescaled Competing Species Processes to a Class of SPDEs.

on $\mathbb{Z}$, i.e. $0 \le p(x) \le 1$ and $\sum_{x\in\mathbb{Z}} p(x) = 1$ such that

$$0 \to 1 \text{ at rate } c(x,\xi) = \sum_y p(x-y)\xi(y),$$
$$1 \to 0 \text{ at rate } c(x,\xi) = \sum_y p(x-y)(1-\xi(y)).$$

Under certain conditions on the rate or kernel, it can now be shown that the given rates determine indeed a unique, $\{0,1\}^{\mathbb{Z}}$-valued Markov process ξ_t.

A possible interpretation of the kernel $p(\cdot)$ is that at exponential times with rate 1, the individual $x \in \mathbb{Z}$ selects a site at random according to the kernel $p(x-\cdot)$ and in case this site has opposite opinion, changes its opinion to the opinion of the selected site. The exponential times and choices according to the random kernel are independent for all $x \in \mathbb{Z}$.

Finally observe that a special case of this model is the case where we fix a finite set $\mathcal{N} \subset \mathbb{Z}$ of neighbours of 0. If we choose the random walk kernel $p(x) = \frac{1}{|\mathcal{N}|}1(x \in \mathcal{N})$, then a neighbour gets chosen with equal probability. Moreover,

$$f_i(x,\xi) = \frac{1}{|\mathcal{N}|}\sum_{y\in\mathcal{N}} 1(\xi(y)=i) = \sum_{y\in\mathbb{Z}} p(x-y)1(\xi(y)=i), \quad i = 0,1 \tag{4.1}$$

can be understood as the frequency of type i in the neighbourhood $x+\mathcal{N}$ of x in configuration ξ.

In general, we can set $f_i(x,\xi) = \sum_{y\in\mathbb{Z}} p(x-y)1(\xi(y)=i), i = 0,1$ to rewrite the rates from above to

$$\begin{aligned} &0 \to 1 \text{ at rate } c(x,\xi) = f_1(x,\xi), \\ &1 \to 0 \text{ at rate } c(x,\xi) = f_0(x,\xi). \end{aligned} \tag{4.2}$$

For the Lotka-Volterra model we consider rate-changes

$$\begin{aligned} &0 \to 1 \text{ at rate} \\ &c(x,\xi) = f_1(x,\xi)\left(f_0(x,\xi) + \alpha_{01} f_1(x,\xi)\right) = f_1(x,\xi) + (\alpha_{01}-1)f_1^2(x,\xi), \\ &1 \to 0 \text{ at rate} \\ &c(x,\xi) = f_0(x,\xi)\left(f_1(x,\xi) + \alpha_{10} f_0(x,\xi)\right) = f_0(x,\xi) + (\alpha_{10}-1)f_0^2(x,\xi) \end{aligned} \tag{4.3}$$

instead, where we used that $f_0(x,\xi) + f_1(x,\xi) = 1$ by definition. The definition will become clear in the Subsection to follow (choose $\lambda = 1$). Observe in particular that if we choose α_{01}, α_{10} close to 1, the Lotka-Volterra model can be seen as a small perturbation of the voter model.

Finally, we can consider biased voter models by multiplying the rate $c(x,\xi)$ of the change $0 \to 1$ by a factor of $(1+\tau)$, i.e. (4.2) becomes

$$\begin{aligned} &0 \to 1 \text{ at rate } c(x,\xi) = (1+\tau) f_1(x,\xi), \\ &1 \to 0 \text{ at rate } c(x,\xi) = f_0(x,\xi) \end{aligned} \tag{4.4}$$

instead. For $\tau > 0$ small we thus have a slight favour for type 1 and for $\tau < 0$ small we have a slight favour for type 0. The biased Lotka-Volterra model is constructed analogously.

4.1.2 Spatial versions of the Lotka-Volterra model

As a further example consider spatial versions of the Lotka-Volterra model with finite range as introduced in [10] (they considered $\xi(x) \in \{1, 2\}$ instead of $\{0, 1\}$). They use rates

$$\begin{aligned} 0 \to 1 \text{ at rate } c(x, \xi) &= \frac{\lambda f_1(x, \xi)}{\lambda f_1(x, \xi) + f_0(x, \xi)} \left(f_0(x, \xi) + \alpha_{01} f_1(x, \xi) \right), \qquad (4.5) \\ 1 \to 0 \text{ at rate } c(x, \xi) &= \frac{f_0(x, \xi)}{\lambda f_1(x, \xi) + f_0(x, \xi)} \left(f_1(x, \xi) + \alpha_{10} f_0(x, \xi) \right), \end{aligned}$$

where $\alpha_{01}, \alpha_{10} \geq 0, \lambda > 0$. Here f_i is as in (4.1) and $\mathcal{N} = \{y : 0 < |y| \leq R\}$ with $R \geq 1$.

We can think of R as the finite interaction range of the model. [10] use this model to obtain results on the parameter regions for coexistence, founder control and spatial segregation of types 0 and 1 in the context of a model that incorporates short-range interactions and dispersal. As a conclusion they obtain that the short-range interactions alter the predictions of the mean-field model.

Following [10] we can interpret the rates as follows. The second multiplicative factor of the rate governs the density-dependent mortality of a particle, the first factor represents the strength of the instantaneous replacement by a particle of opposite type. The mortality of type 0 consists of two parts, f_0 describes the effect of intraspecific competition, $\alpha_{01} f_1$ the effect of interspecific competition. [10] assume that the intraspecific competition is the same for both species. The replacement of a particle of opposite type is regulated by the fecundity parameter λ. The first factors of both rate-changes added together yield 1. Thus they can be seen as weighted densities of the two species. If $\lambda > 1$, species 1 has a higher fecundity than species 0.

4.1.3 Long-range limits

In [9], Mueller and Tribe show that the approximate densities of type 1 of rescaled biased voter processes, defined as in (4.1) and (4.4) with $\tau = \frac{\theta}{N}$, converge to continuous space time densities which solve the heat equation with drift, driven by Fisher-Wright noise, namely

$$\frac{\partial u}{\partial t} = \frac{\Delta u}{6} + 2\theta(1-u)u + \sqrt{4u(1-u)}\dot{W}. \qquad (4.6)$$

Observe that [9] scale space by $1/N$ and consider $\mathcal{N} = \{0 < |y| \leq N^{-1/2}\}$. Hence, the number of neighbours of $x \in \mathbb{Z}/N$ is increasing, namely $|\mathcal{N}| = 2c(N)N^{1/2}$ with $c(N) \overset{N \to \infty}{\to} 1$, and we thus obtain long-range interactions. Finally, they also rescale time by speeding up the rates of change $c(x, \xi)$ as follows.

Let $p^{(N)}(x) = \frac{1}{|\mathcal{N}|}1(x \in \mathcal{N})$, then in the N^{th}-model we have

$$\begin{aligned} 0 \to 1 \text{ at rate } c^N\big(x,\xi^N\big) &= \Big(N^{1/2} + \theta N^{-1/2}\Big)\,|\mathcal{N}| \sum_{y\in\mathbb{Z}/N} p^{(N)}(x-y)\xi^N(y) \\ &= 2c(N)(N+\theta)f_1\big(x,\xi^N\big)\,, \\ 1 \to 0 \text{ at rate } c^N\big(x,\xi^N\big) &= N^{1/2}|\mathcal{N}| \sum_{y\in\mathbb{Z}/N} p^{(N)}(x-y)(1-\xi^N(y)) \\ &= 2c(N)Nf_0\big(x,\xi^N\big)\,. \end{aligned}$$

They fix $\theta \geq 0$, i.e. consider the case where the opinion of type 1 is slightly dominant. See the introduction and Theorem 2 of [9] for more details.

In [3] it was shown that stochastic spatial Lotka-Volterra models, suitably rescaled in space and time, converge weakly to super-Brownian motion with linear drift. As they choose the parameters α_{01}, α_{10} (recall (4.3)) such that

$$N\left(\alpha_{i(1-i)}^{(N)} - 1\right) \to \theta_i \in \mathbb{R}, \quad i = 0, 1 \tag{4.7}$$

(see (H3) in [3]) their models can also be interpreted as small perturbations of the voter model. [3] extended the main results of Cox, Durrett and Perkins [2], which proved similar results for long-range voter models. Both papers treat the low density regime, i.e. where only a finite number of individuals of type 1 is present. Instead of investigating limits for approximate densities, both papers define measure-valued processes X_t^N by

$$X_t^N = \frac{1}{N'} \sum_{x\in\mathbb{Z}/(M_N\sqrt{N})} \xi_t^N(x)\delta_x,$$

i.e. they assign mass $1/N'$, $N' = N'(N)$ to each individual of type 1 and consider weak limits in the space of finite Borel measures on $\mathbb{R}$. In particular, they establish the tightness of the sequence of measures and the uniqueness of the martingale problem, solved by any limit point.

Note that both papers use a different scaling in comparison to [9]. Using the notation in [2], for $d = 1$ they take $N' = N$ and the space is scaled by $M_N\sqrt{N}$ with $M_N/\sqrt{N} \to \infty$ (see for instance Theorem 1.1 of [2] for $d = 1$) in the long-range setup. According to this notation, [9] used $M_N = \sqrt{N}$, which is at the threshold of the results in [2], but not included. By letting $M_N = \sqrt{N}$ in our setup several non-linear terms will arise in our limiting SPDE below. Also note the brief discussion of the case where $M_N/\sqrt{N} \to 0$ in $d = 1$ before (H3) in [2].

Additionally, [2] and [3] consider fixed kernel models in dimensions $d \geq 2$ respectively $d \geq 3$ with $M_N = 1$ and a fixed random walk kernel q satisfying some additional conditions such that $p(x) = q(\sqrt{N}x)$ on $x \in \mathbb{Z}/(M_N\sqrt{N})$.

Finally, in Cox and Perkins [4], the results of [3] for $d \geq 3$ are used to relate the limiting super-Brownian motions to questions of coexistence and survival of a rare type in the original Lotka-Volterra model.

4.1.4 Overview of results

In the present paper we first prove tightness of the local densities for scaling limits of more general particle systems. The generalization includes two features.

Firstly, we shall extend the model in [9] to limits of small perturbations of the long-range voter model, including the setup from [10]. As the rates in [10] (see (4.5)) include taking ratios, we extend our perturbations to a set of power series (for extensions to polynomials of degree 2 recall (4.3)), thereby including certain analytic functions. Recall in particular from (4.7) that we shall allow the coefficients of the power series to depend on N.

Secondly, we shall combine both long-range interaction and fixed kernel interaction for the perturbations. As we shall see, the tightness results will carry over. As a special case we shall be able to consider rescaled Lotka-Volterra models with long-range dispersal and short-range competition, i.e. where (4.3) gets generalized to

$$\begin{aligned} 0 \to 1 \text{ at rate } c(x,\xi) &= f_1(x,\xi)\left(g_0(x,\xi) + \alpha_{01} g_1(x,\xi)\right), \\ 1 \to 0 \text{ at rate } c(x,\xi) &= f_0(x,\xi)\left(g_1(x,\xi) + \alpha_{10} g_0(x,\xi)\right). \end{aligned}$$

Here $f_i(x,\xi), i = 0,1$ is the density corresponding to a long-range kernel p^L and $g_i(x,\xi), i = 0,1$ is the density corresponding to a fixed kernel p^F (also recall the interpretation of both multiplicative factors in Subsection 4.1.2).

Finally, in the case of long-range interactions only we show the limit points are solutions of a SPDE similar to (4.6) but with a drift depending on the choice of our perturbation and small changes in constants due to simple differences in scale factors. Hence, we obtain a class of SPDEs that can be characterized as the limit of perturbations of the long-range voter model. If the limiting initial condition u_0 satisfies $\int u_0(x)dx < \infty$, we can show the weak uniqueness of solutions to the limiting SPDE and therefore show weak convergence of the rescaled particle densities to this unique law.

When there exists a fixed kernel, the question of uniqueness of all limit points and of identifying the limit remains an open problem. Also, when we consider long-range interactions only with $\int u_0(x)dx = \infty$ the proof of weak uniqueness of solutions to the limiting SPDE remains open.

The proof of our results generalizes the work done in [9]. In [9], limits are considered for both the long-range contact process and the long-range voter process. Full details are given for the contact process. For the voter process, once the approximate martingale problem is derived, almost all of the remaining steps are left to the reader. Many arguments of our proof are similar to [9] but as additions and adaptations are needed due to our broader setup and as they did not provide details for the long-range voter model we shall sometimes be more detailed.

4.1.5 Outline of the paper

In Section 4.2 to follow we shall first set up our model and give the main results. Then we shall reformulate the model so that it can be approached by the methods used in [9]. A statement of the main results in the reformulated setting follows.

In Section 4.3 we shall introduce a graphical construction for each approximating model $\xi_\cdot^N$. This allows us to write out the time-evolution of our models. By integrating it against a test function and summing over $x \in \mathbb{Z}/N$ we finally obtain an approximate martingale problem for the N^{th}-process. We define the approximate density $A(\xi_t^N)(x)$ as the average density of particles of type 1 on $\mathbb{Z}/N$ in an interval centered at x of length $2/\sqrt{N}$ (see (4.10) below). By choosing a specific test function, the properties of which are under investigation at the beginning of Section 4.4, an approximate Green's function representation for the approximate densities $A(\xi_t^N)(\cdot)$ is derived towards the end of Section 4.4 and bounds on error-terms appearing in it are given. Making use of the Green's function representation, tightness of $A(\xi_t^N)(\cdot)$ is proven in Section 4.5. Here the main part of the proof consists in finding estimates on p^{th}-moment differences.

In Section 4.6 the tightness of the approximate densities is used to show tightness of the measure corresponding to the sequence of configurations ξ_t^N. Finally, in the special case with no fixed kernel, every limit is shown to solve a certain SPDE. In Section 4.7 we prove that this SPDE has a unique weak solution if $< u_0, 1 >< \infty$. In this case, weak uniqueness of the limits of the sequence of approximate densities follows.

4.2 Main Results of the Paper

4.2.1 The model

We define a sequence of rescaled competing species models ξ_t^N in dimension $d = 1$, which can be described as perturbations of voter models. In the N^{th}-model the sites are indexed by $x \in N^{-1}\mathbb{Z}$. We label the state of site x at time t by $\xi_t^N(x)$ where $\xi_t^N(x) = 0$ if the site is occupied at time t by type 0 and $\xi_t^N(x) = 1$ if it is occupied by type 1.

In what follows we shall write $x \sim y$ if and only if $0 < |x - y| \leq N^{-1/2}$, i.e. if and only if x is a neighbour of y. Observe that each x has

$$2c(N)N^{1/2}, \qquad c(N) \overset{N\to\infty}{\to} 1$$

neighbours.

The rates of change incorporate both long-range models and fixed kernel models with finite range. The long-range interaction takes into account the densities of the neighbours of x at long-range, i.e.

$$f_i^{(N)}(x, \xi) \equiv \frac{1}{2c(N)\sqrt{N}} \sum_{\substack{0<|y-x|\leq 1/\sqrt{N}, \\ y\in\mathbb{Z}/N}} 1(\xi^N(y) = i), \quad i = 0, 1$$

and the fixed kernel interaction considers

$$g_i^{(N)}(x,\xi) \equiv \sum_{y\in\mathbb{Z}/N} p(N(x-y))1(\xi^N(y)=i), \quad i=0,1,$$

where $p(x)$ is a random walk kernel on $\mathbb{Z}$ of finite range, i.e. $0 \le p(x) \le 1$, $\sum_{x\in\mathbb{Z}} p(x) = 1$ and $p(x) = 0$ for all $|x| \ge C_p$. In what follows we shall often abbreviate $f_i^{(N)}(x,\xi)$ by $f_i^{(N)}$ and $g_i^{(N)}(x,\xi)$ by $g_i^{(N)}$ if the context is clear.

Now define the rates of change of our configurations. At site x in configuration $\xi^N \in \{0,1\}^{\mathbb{Z}/N}$ the coordinate $\xi^N(x)$ makes transitions

$$\begin{aligned} &0 \to 1 \text{ at rate } Nf_1^{(N)} + f_1^{(N)}\left\{g_0^{(N)}G_0^{(N)}\left(f_1^{(N)}\right) + g_1^{(N)}H_0^{(N)}\left(f_1^{(N)}\right)\right\}, \quad (4.8)\\ &1 \to 0 \text{ at rate } Nf_0^{(N)} + f_0^{(N)}\left\{g_0^{(N)}G_1^{(N)}\left(f_0^{(N)}\right) + g_1^{(N)}H_1^{(N)}\left(f_0^{(N)}\right)\right\}, \end{aligned}$$

where $G_i^{(N)}, H_i^{(N)}, i=0,1$ are power series as follows.

Hypothesis 4.2.1. *We assume that*

$$G_i^{(N)}(x) \equiv \sum_{m=0}^{\infty} \alpha_i^{(m+1,N)} x^m \text{ and } H_i^{(N)}(x) \equiv \sum_{m=0}^{\infty} \beta_i^{(m+1,N)} x^m, \; x \in [0,1]$$

with $i=0,1$, $\alpha_i^{(m+1,N)}, \beta_i^{(m+1,N)} \in \mathbb{R}$, $m \ge 0$ *and that there exists* $N_0 \in \mathbb{N}$ *such that*

$$\begin{aligned} &\sup_{N\ge N_0} \sum_{i=0,1} \sum_{m=0}^{\infty} \left\{\alpha_i^{(m+1,N)} \vee 0 + \beta_i^{(m+1,N)} \vee 0 + m\left|\alpha_i^{(m+1,N)} \wedge 0\right|\right.\\ &\qquad \left. +m\left|\beta_i^{(m+1,N)} \wedge 0\right|\right\} < \infty. \end{aligned}$$

Remark 4.2.2. *The above rates determine indeed a unique,* $\{0,1\}^{\mathbb{Z}/N}$*-valued Markov process* ξ_t^N *for* $N \ge N_0$ *with* N_0 *as in Hypothesis 4.2.1 as we now show. See for instance Theorem B3, p.3 in Liggett [8] and note the uniform boundedness assumption on the rates from p.1 of [8]. Following the notation in [8], let* $c(x,\xi^N)$ *denote the rate at which the coordinate* $\xi^N(x)$ *flips from* 0 *to* 1 *or from* 1 *to* 0 *when the system is in state* ξ^N. *Then using (4.8),* $0 \le f_i^{(N)}, g_i^{(N)} \le 1, i=0,1$ *and Hypothesis 4.2.1 yields*

$$\begin{aligned} &\sup_{x\in\mathbb{Z}/N} \sup_{\xi^N\in\{0,1\}^{\mathbb{Z}/N}} c\big(x,\xi^N\big)\\ &\le N + \left(\sum_{m=0}^{\infty}\left|\alpha_0^{(m+1,N)}\right| + \left|\beta_0^{(m+1,N)}\right|\right) \vee \left(\sum_{m=0}^{\infty}\left|\alpha_1^{(m+1,N)}\right| + \left|\beta_1^{(m+1,N)}\right|\right)\\ &\equiv N + C_0(N) < \infty \end{aligned}$$

and

$$\begin{aligned}
&\sup_{x\in\mathbb{Z}/N}\sum_{u\in\mathbb{Z}/N}\sup_{\xi^N\in\{0,1\}^{\mathbb{Z}/N}}\left|c\big(x,\xi^N\big)-c\big(x,\xi^N_u\big)\right|\\
&\le \sup_{x\in\mathbb{Z}/N}\sum_{u\sim x}\sup_{\xi^N\in\{0,1\}^{\mathbb{Z}/N}}\left|c\big(x,\xi^N\big)-c\big(x,\xi^N_u\big)\right|\\
&\quad+\sup_{x\in\mathbb{Z}/N}\sum_{u\in\mathbb{Z}/N}\sup_{\xi^N\in\{0,1\}^{\mathbb{Z}/N}}\sum_{i=0,1}\left|g_i^{(N)}\big(x,\xi^N\big)-g_i^{(N)}\big(x,\xi^N_u\big)\right|C_0(N)\\
&\le 2c(N)N^{1/2}2(N+C_0(N))+\sup_{x\in\mathbb{Z}/N}\sum_{u\in\mathbb{Z}/N}2p(N(x-u))C_0(N)\\
&\le 2c(N)N^{1/2}2(N+C_0(N))+2C_0(N)\\
&<\infty,
\end{aligned}$$

where

$$\xi^N_u(v)=\begin{cases}\xi^N(v), & v\neq u,\\ 1-\xi^N(v), & v=u.\end{cases}$$

Following [8], the two conditions are sufficient to ensure the above.

Additionally, the closure in the space of continuous functions on $\{0,1\}^{\mathbb{Z}/N}$ *of the operator* $\Omega f\big(\xi^N\big)=\sum_x c\big(x,\xi^N\big)\left(f(\xi^N_x)-f(\xi^N)\right)$*, which is defined on the space of finite cylinder functions on* $\{0,1\}^{\mathbb{Z}/N}$*, is the Markov generator of the process* ξ^N_t*.*

Remark 4.2.3. *Observe in particular that* $f_0^{(N)}+f_1^{(N)}=1$ *and* $g_0^{(N)}+g_1^{(N)}=1$*. Hence the special case of no fixed kernel can be obtained by choosing* $G_i^{(N)}\equiv H_i^{(N)}, i=0,1$ *and we get*

$$\begin{aligned}
&0\to 1 \text{ at rate } Nf_1^{(N)}+f_1^{(N)}G_0^{(N)}\left(f_1^{(N)}\right),\\
&1\to 0 \text{ at rate } Nf_0^{(N)}+f_0^{(N)}G_1^{(N)}\left(f_0^{(N)}\right).
\end{aligned}\tag{4.9}$$

For the configurations $\xi^N_t\in\{0,1\}^{\mathbb{Z}/N}$ we define approximate densities $A(\xi^N_t)$ via

$$A(\xi^N_t)(x)=\frac{1}{2c(N)N^{1/2}}\sum_{y\sim x}\xi^N_t(y),\qquad x\in N^{-1}\mathbb{Z}\tag{4.10}$$

and note that $A(\xi^N_t)(x)=f_1^{(N)}\big(x,\xi^N_t\big)$. By linearly interpolating between sites we obtain approximate densities $A(\xi^N_t)(x)$ for all $x\in\mathbb{R}$.

Notation 4.2.4. Set $\mathcal{C}_1\equiv\{f:\mathbb{R}\to[0,1]\text{ cont.}\}$ and let $\mathcal{C}_1$ be equipped with the topology of uniform convergence on compact sets.

We obtain that $t\mapsto A(\xi^N_t)$ is cadlag $\mathcal{C}_1$-valued, where we used that

$$0\le A(\xi^N_t)(x)\le 1 \text{ for all } x\in N^{-1}\mathbb{Z}.$$

Hence, we can consider the law of $A(\xi^N)$ on the space of cadlag $\mathcal{C}_1$-valued paths with the Skorokhod topology.

4.2.2 Main results

Before stating our main results we need some more notation.

Notation 4.2.5. For $f, g : N^{-1}\mathbb{Z} \to \mathbb{R}$, we set

$$< f, g >= \frac{1}{N} \sum_x f(x)g(x).$$

Let ν be a measure on $N^{-1}\mathbb{Z}$. Then we set

$$< \nu, f >= \int f d\nu.$$

Remark 4.2.6. *We can rewrite every configuration ξ_t^N in terms of its corresponding measure. Let*

$$\nu_t^N \equiv \frac{1}{N} \sum_x \delta_x 1(\xi_t^N(x) = 1),$$

then

$$< \xi_t^N, \phi >=< \nu_t^N, \phi > .$$

Definition 4.2.7. Let S be a Polish space and let $D(S)$ denote the space of cadlag paths from $\mathbb{R}_+$ to S with the Skorokhod topology. Following the first definition on p.148 of Perkins [11], we shall say that a collection of processes with paths in $D(S)$ is C-tight if and only if it is tight in $D(S)$ and all weak limit points are a.s. continuous. Recall that for Polish spaces, tightness and weak relative compactness are equivalent.

Remark 4.2.8. *In what follows we shall investigate tightness of $\{A(\xi_\cdot^N) : N \geq N_0\}$ in $D(\mathcal{C}_1)$ and tightness of $\{\nu_t^N : N \geq N_0\}$ in $D(\mathcal{M}(\mathbb{R}))$, where $\mathcal{M}(\mathbb{R})$ is the space of Radon measures equipped with the vague topology ($\mathcal{M}(\mathbb{R})$ is indeed Polish, see Dawson [5], Section 3.1.3).*

Theorem 4.2.9. *Suppose that $A(\xi_0^N) \to u_0$ in $\mathcal{C}_1$. Let the transition rates of $\xi^N(x)$ be as in (4.8) with $G_i^{(N)}, H_i^{(N)}, i = 0, 1$ satisfying Hypothesis 4.2.1. Then $\left(A(\xi_t^N) : t \geq 0\right)$ are C-tight as cadlag $\mathcal{C}_1$-valued processes and the $\left(\nu_t^N : t \geq 0\right)$ are C-tight as cadlag Radon measure valued processes with the vague topology. If $\left(A\left(\xi_t^{N_k}\right), \nu_t^{N_k}\right)_{t\geq 0}$ converges to $(u_t, \nu_t)_{t\geq 0}$, then $\nu_t(dx) = u_t(x)dx$ for all $t \geq 0$.*

Remark 4.2.10. *The above applies in particular to models where $G_i^{(N)}, H_i^{(N)}$ are finite sums (see Hypothesis 4.2.1).*

Hypothesis 4.2.11. *Let us consider the special case with no fixed kernel (see Remark 4.2.3). Additionally to Hypothesis 4.2.1 we assume that*

$$\alpha_i^{(m+1,N)} \overset{N\to\infty}{\to} \alpha_i^{(m+1)} \text{ for all } i = 0, 1 \text{ and } m \geq 0$$

with (take sgn$(0)=0$*)*

$$sgn\left(\alpha_i^{(m+1,N)}\right) \geq 0 \text{ for all } N \geq N_0 \text{ or } sgn\left(\alpha_i^{(m+1,N)}\right) \leq 0 \text{ for all } N \geq N_0$$

for all $i = 0, 1, m \geq 0$ *and that*

$$\lim_{N\to\infty} \sum_{i=0,1} \sum_{m=0}^{\infty} \left\{ \alpha_i^{(m+1,N)} \vee 0 + m \left| \alpha_i^{(m+1,N)} \wedge 0 \right| \right\} \tag{4.11}$$
$$= \sum_{i=0,1} \sum_{m=0}^{\infty} \left\{ \alpha_i^{(m+1)} \vee 0 + m \left| \alpha_i^{(m+1)} \wedge 0 \right| \right\}.$$

Remark 4.2.12. *The additional conditions of Hypothesis 4.2.11 are necessary to transform the given rates into rates with positive coefficients in a uniform way in Subsection 4.2.3 and to later characterize limit points of the approximate densities by taking limits in* $N \to \infty$ *inside infinte sums in (4.55).*

Definition 4.2.13. Under the assumptions of Theorem 4.2.9 and Hypothesis 4.2.11 we let for $x \in [0,1]$,

$$G_i(x) \equiv \lim_{N\to\infty} G_i^{(N)}(x) = \lim_{N\to\infty} \sum_{m=0}^{\infty} \alpha_i^{(m+1,N)} x^m = \sum_{m=0}^{\infty} \alpha_i^{(m+1)} x^m, i = 0, 1.$$

This is well-defined by (4.11) and Royden [12], Proposition 11.18.

Theorem 4.2.14. *We obtain under the assumptions of Theorem 4.2.9 and Hypothesis 4.2.11 for the special case with no fixed kernel that the limit points of* $A(\xi_t^N)$ *are continuous* $\mathcal{C}_1$*-valued processes* u_t *which solve*

$$\frac{\partial u}{\partial t} = \frac{\Delta u}{6} + (1-u)u\left\{G_0(u) - G_1(1-u)\right\} + \sqrt{2u(1-u)}\dot{W} \tag{4.12}$$

with initial condition u_0*. If we assume additionally* $< u_0, 1 >< \infty$*, then* u_t *is the unique (in law)* $[0,1]$*-valued solution to the above SPDE.*

Remark 4.2.15. *As an example consider spatial versions of the Lotka-Volterra model as introduced in Subsection 4.1.2. In what follows we shall choose the competition and fecundity parameters near one and we shall consider the long-range case. Namely, the model exhibits the following rates:*

$$0 \to 1 \text{ at rate } N\left[\frac{\lambda^{(N)} f_1^{(N)}}{\lambda^{(N)} f_1^{(N)} + f_0^{(N)}}\left(f_0^{(N)} + \alpha_{01}^{(N)} f_1^{(N)}\right)\right],$$
$$1 \to 0 \text{ at rate } N\left[\frac{f_0^{(N)}}{\lambda^{(N)} f_1^{(N)} + f_0^{(N)}}\left(f_1^{(N)} + \alpha_{10}^{(N)} f_0^{(N)}\right)\right].$$

We suppose that

$$\lambda^{(N)} \equiv 1 + \frac{\lambda'}{N},\ \alpha_{01}^{(N)} \equiv 1 + \frac{\alpha_{01}}{N},\ \alpha_{10}^{(N)} \equiv 1 + \frac{\alpha_{10}}{N}.$$

Using $f_0^{(N)} + f_1^{(N)} = 1$ *we can therefore rewrite the rates as*

$$0 \to 1 \text{ at rate } (N+\lambda') f_1^{(N)} \left(1 + \frac{\alpha_{01}}{N} f_1^{(N)}\right) \sum_{n\geq 0} \left(-\frac{\lambda'}{N} f_1^{(N)}\right)^n, \tag{4.13}$$

$$\begin{aligned} 1 \to 0 \text{ at rate } & N f_0^{(N)} \left(1 + \frac{\alpha_{10}}{N} f_0^{(N)}\right) \sum_{n\geq 0} \left(-\frac{\lambda'}{N} f_1^{(N)}\right)^n \\ &= N f_0^{(N)} \left(1 + \frac{\alpha_{10}}{N} f_0^{(N)}\right) \sum_{k\geq 0} \left(f_0^{(N)}\right)^k \left(\frac{\lambda'}{N}\right)^k \sum_{n\geq k} \binom{n}{k} \left(-\frac{\lambda'}{N}\right)^{n-k}. \end{aligned}$$

Here we used that $\left|f_i^{(N)}\right| \leq 1, i = 0,1$ *and that* $\left|\frac{\lambda'}{N}\right| \to 0$ *for* $N \to \infty$. *We can use the explicit calculations for a geometric series, in particular that we have* $\sum_{n\geq 0} n|q|^n < \infty$ *and* $\sum_{n\geq k} |q|^{n-k}\binom{n}{k} = \frac{1}{(1-|q|)^{k+1}}$ *for* $|q| < 1$ *to check that Hypothesis 4.2.1 and Hypothesis 4.2.11 are satisfied.*

Using Theorem 4.2.14 we further obtain that the limit points of $A(\xi_t^N)$ *are continuous* $\mathcal{C}_1$*-valued processes* u_t *which solve*

$$\begin{aligned} \frac{\partial u}{\partial t} =& \frac{\Delta u}{6} + (1-u)u\left\{(\lambda' + u(\alpha_{01} - \lambda')) - (-\lambda' + (1-u)(\alpha_{10} + \lambda'))\right\} \\ &+ \sqrt{2u(1-u)}\dot{W} \\ =& \frac{\Delta u}{6} + (1-u)u\left\{\lambda' - \alpha_{10} + u(\alpha_{01} + \alpha_{10})\right\} + \sqrt{2u(1-u)}\dot{W} \end{aligned}$$

by rewriting the above rates (4.13) in the form (4.9) and taking the limit for $N \to \infty$. *For* $< u_0, 1 >< \infty$, u_t *is the unique weak* $[0,1]$*-valued solution to the above SPDE.*

4.2.3 Reformulation

We proceed as in [9]. Recall from Hypothesis 4.2.1 and (4.8) that the rates of change are given by

$$0 \to 1 \text{ at rate } N f_1^{(N)} + g_0^{(N)} \sum_{m=1}^{\infty} \alpha_0^{(m,N)} \left(f_1^{(N)}\right)^m + g_1^{(N)} \sum_{m=1}^{\infty} \beta_0^{(m,N)} \left(f_1^{(N)}\right)^m, \tag{4.14}$$

$$1 \to 0 \text{ at rate } N f_0^{(N)} + g_0^{(N)} \sum_{m=1}^{\infty} \alpha_1^{(m,N)} \left(f_0^{(N)}\right)^m + g_1^{(N)} \sum_{m=1}^{\infty} \beta_1^{(m,N)} \left(f_0^{(N)}\right)^m,$$

where $\alpha_j^{(m,N)}, \beta_j^{(m,N)} \in \mathbb{R}$ for all $j = 0,1, m \in \mathbb{N}$.

Following [9] we shall model each term in the rate-changes via independent families of i.i.d. Poisson processes. For instance, if $\alpha_0^{(m,N)}$ is non-negative, the term $g_0^{(N)} \alpha_0^{(m,N)} \left(f_1^{(N)}\right)^m$ of the rate-change $0 \to 1$ in (4.14) can be modeled via i.i.d. Poisson processes

$$\left(Q_t(x; y_1, \ldots, y_m; z) : x, y_1, \ldots, y_m, z \in N^{-1}\mathbb{Z}\right)$$

of rate

$$\frac{\alpha_0^{(m,N)}}{(2c(N))^m N^{m/2}} p(N(x-z)).$$

At a jump of $Q_t(x;y_1,\ldots,y_m;z)$ the voter at x adopts the opinion 1 provided that all of $y_1,\ldots,y_m$ have opinion 1 and z has opinion 0.

As we want to allow the $\alpha_i^{(m,N)}, \beta_i^{(m,N)}$ to be negative, too, we first rewrite (4.14) with the help of $f_0^{(N)} + f_1^{(N)} = 1$ and $g_0^{(N)} + g_1^{(N)} = 1$ in a form where all resulting coefficients are non-negative.

Corollary 4.2.16. *We can rewrite our transitions as follows.*

$0 \to 1$ *at rate* (4.15)

$$\left(N - \theta^{(N)}\right) f_1^{(N)} + f_1^{(N)} \left\{ \sum_{i=0,1} a_i^{(N)} g_i^{(N)} + \sum_{m\geq 2, i,j=0,1} q_{ij}^{(0,m,N)} g_i^{(N)} f_j^{(N)} \left(f_1^{(N)}\right)^{m-2} \right\},$$

$1 \to 0$ *at rate*

$$\left(N - \theta^{(N)}\right) f_0^{(N)} + f_0^{(N)} \left\{ \sum_{i=0,1} b_i^{(N)} g_i^{(N)} + \sum_{m\geq 2, i,j=0,1} q_{ij}^{(1,m,N)} g_i^{(N)} f_j^{(N)} \left(f_0^{(N)}\right)^{m-2} \right\},$$

with corresponding $\theta^{(N)}, a_i^{(N)}, b_i^{(N)}, q_{ij}^{(k,m,N)} \in \mathbb{R}^+, i,j,k = 0,1, m \geq 2.$

Proof. We shall drop the superscripts of $f_i^{(N)}, g_i^{(N)}, i = 0,1$ in what follows to simplify notation.

Suppose for instance $\alpha_0^{(m,N)} < 0$ for some $m \geq 1$ in (4.14). Using that

$$-x^m = (1-x)\sum_{k=1}^{m-1} x^k - x$$

and recalling that $1 - f_1 = f_0$ we obtain

$$g_0 \alpha_0^{(2k+1,N)} f_1^{2k+1} = g_0 \left\{ \left(-\alpha_0^{(2k+1,N)}\right) f_0 \sum_{l=1}^{2k} f_1^l + \alpha_0^{(2k+1,N)} f_1 \right\}.$$

Finally, we can use $g_0 = 1 - g_1$ to obtain

$$g_0 \alpha_0^{(2k+1,N)} f_1^{2k+1} = g_0 \left(-\alpha_0^{(2k+1,N)}\right) f_0 \sum_{l=1}^{2k} f_1^l + g_1 \left(-\alpha_0^{(2k+1,N)}\right) f_1 + \alpha_0^{(2k+1,N)} f_1.$$

All terms on the r.h.s. but the last can be accommodated into an existing representation (4.15) as follows:

$$\begin{aligned} q_{00}^{(0,m,N)} &\to q_{00}^{(0,m,N)} + \left(-\alpha_0^{(2k+1,N)}\right) \text{ for } 2 \leq m \leq 2k+1, \\ a_1^{(N)} &\to a_1^{(N)} + \left(-\alpha_0^{(2k+1,N)}\right). \end{aligned}$$

Finally, we can assimilate the last term into the first part of the rate $0 \to 1$, i.e. we replace

$$\theta^{(N)} \to \theta^{(N)} - \alpha_0^{(2k+1,N)}.$$

As we use the representation (4.15), a change in the first part of the rate $0 \to 1$ also impacts the rate $1 \to 0$ in its first term. Therefore we have to fix the rate $1 \to 0$ by adding a term of $(-\alpha_0^{(2k+1,N)})f_0 = g_0 f_0(-\alpha_0^{(2k+1,N)}) + g_1 f_0(-\alpha_0^{(2k+1,N)})$ to the second and third term of the rate, i.e. by replacing

$$b_0^{(N)} \to b_0^{(N)} + \left(-\alpha_0^{(2k+1,N)}\right),$$
$$b_1^{(N)} \to b_1^{(N)} + \left(-\alpha_0^{(2k+1,N)}\right).$$

The case $m = 2k, k \geq 1$ follows similarly and the general case with multiple negative $\alpha's$ and/or $\beta's$ now follows inductively. ☐

Remark 4.2.17. *The above construction yields the following non-negative coefficients:*

$$\theta^{(N)} \equiv \sum_{j=0,1} \sum_{n=1}^{\infty} \left(-\alpha_j^{(n,N)}\right) 1\left(\alpha_j^{(n,N)} < 0\right) + \left(-\beta_j^{(n,N)}\right) 1\left(\beta_j^{(n,N)} < 0\right),$$

$$q_{00}^{(0,m,N)} \equiv \sum_{n=m}^{\infty} \left(-\alpha_0^{(n,N)}\right) 1\left(\alpha_0^{(n,N)} < 0\right),$$

$$a_0^{(N)} \equiv \alpha_0^{(1,N)} 1\left(\alpha_0^{(1,N)} \geq 0\right) + \left\{\sum_{n=1}^{\infty} \left(-\beta_0^{(n,N)}\right) 1\left(\beta_0^{(n,N)} < 0\right) + \sum_{n=1}^{\infty} \left(-\alpha_1^{(n,N)}\right) 1\left(\alpha_1^{(n,N)} < 0\right) + \sum_{n=1}^{\infty} \left(-\beta_1^{(n,N)}\right) 1\left(\beta_1^{(n,N)} < 0\right)\right\},$$

$$q_{01}^{(0,m,N)} \equiv \alpha_0^{(m,N)} 1\left(\alpha_0^{(m,N)} \geq 0\right), \qquad q_{10}^{(0,m,N)} \equiv \sum_{n=m}^{\infty} \left(-\beta_0^{(n,N)}\right) 1\left(\beta_0^{(n,N)} < 0\right),$$

$$a_1^{(N)} \equiv \beta_0^{(1,N)} 1\left(\beta_0^{(1,N)} \geq 0\right) + \left\{\sum_{n=1}^{\infty} \left(-\alpha_0^{(n,N)}\right) 1\left(\alpha_0^{(n,N)} < 0\right) + \sum_{n=1}^{\infty} \left(-\alpha_1^{(n,N)}\right) 1\left(\alpha_1^{(n,N)} < 0\right) + \sum_{n=1}^{\infty} \left(-\beta_1^{(n,N)}\right) 1\left(\beta_1^{(n,N)} < 0\right)\right\},$$

$$q_{11}^{(0,m,N)} \equiv \beta_0^{(m,N)} 1\left(\beta_0^{(m,N)} \geq 0\right), \qquad q_{01}^{(1,m,N)} \equiv \sum_{n=m}^{\infty} \left(-\alpha_1^{(n,N)}\right) 1\left(\alpha_1^{(n,N)} < 0\right),$$

$$b_0^{(N)} \equiv \alpha_1^{(1,N)} 1\left(\alpha_1^{(1,N)} \geq 0\right) + \left\{\sum_{n=1}^{\infty} \left(-\beta_1^{(n,N)}\right) 1\left(\beta_1^{(n,N)} < 0\right) + \sum_{n=1}^{\infty} \left(-\alpha_0^{(n,N)}\right) 1\left(\alpha_0^{(n,N)} < 0\right) + \sum_{n=1}^{\infty} \left(-\beta_0^{(n,N)}\right) 1\left(\beta_0^{(n,N)} < 0\right)\right\},$$

$$q_{00}^{(1,m,N)} \equiv \alpha_1^{(m,N)} 1\left(\alpha_1^{(m,N)} \geq 0\right), \qquad q_{11}^{(1,m,N)} \equiv \sum_{n=m}^{\infty} \left(-\beta_1^{(n,N)}\right) 1\left(\beta_1^{(n,N)} < 0\right),$$

$$b_1^{(N)} \equiv \beta_1^{(1,N)} 1\left(\beta_1^{(1,N)} \geq 0\right) + \left\{ \sum_{n=1}^{\infty} \left(-\alpha_1^{(n,N)}\right) 1\left(\alpha_1^{(n,N)} < 0\right) \right.$$
$$\left. + \sum_{n=1}^{\infty} \left(-\alpha_0^{(n,N)}\right) 1\left(\alpha_0^{(n,N)} < 0\right) + \sum_{n=1}^{\infty} \left(-\beta_0^{(n,N)}\right) 1\left(\beta_0^{(n,N)} < 0\right) \right\},$$

$$q_{10}^{(1,m,N)} \equiv \beta_1^{(m,N)} 1\left(\beta_1^{(m,N)} \geq 0\right).$$

By Hypothesis 4.2.1 this implies in particular that there exists $N_0 \in \mathbb{N}$ such that

$$\sup_{N \geq N_0} \sum_{i,j,k=0,1} \sum_{m \geq 2} q_{ij}^{(k,m,N)} < \infty.$$

Remark 4.2.18. *Observe that we can rewrite the transition rates in (4.15) such that $a_i^{(N)} = b_i^{(N)} = 0, i = 0, 1,$ i.e.*

$$0 \to 1 \text{ at rate} \tag{4.16}$$
$$\left(N - \theta^{(N)}\right) f_1^{(N)} + f_1^{(N)} \sum_{m \geq 2, i,j=0,1} q_{ij}^{(0,m,N)} g_i^{(N)} f_j^{(N)} \left(f_1^{(N)}\right)^{m-2},$$
$$1 \to 0 \text{ at rate}$$
$$\left(N - \theta^{(N)}\right) f_0^{(N)} + f_0^{(N)} \sum_{m \geq 2, i,j=0,1} q_{ij}^{(1,m,N)} g_i^{(N)} f_j^{(N)} \left(f_0^{(N)}\right)^{m-2}.$$

Indeed, using that $f_0^{(N)} + f_1^{(N)} = 1$ we can change for instance

$$a_0^{(N)} g_0^{(N)} + q_{00}^{(0,2,N)} g_0^{(N)} f_0^{(N)} \left(f_1^{(N)}\right)^0 + q_{01}^{(0,2,N)} g_0^{(N)} f_1^{(N)} \left(f_1^{(N)}\right)^0$$

with $a_0^{(N)}, q_{00}^{(0,2,N)}, q_{01}^{(0,2,N)}$ nonnegative into

$$\left(a_0^{(N)} + q_{00}^{(0,2,N)}\right) g_0^{(N)} f_0^{(N)} \left(f_1^{(N)}\right)^0 + \left(a_0^{(N)} + q_{01}^{(0,2,N)}\right) g_0^{(N)} f_1^{(N)} \left(f_1^{(N)}\right)^0,$$

where the new coefficients are nonnegative again.

Recall Remark 4.2.17 together with Hypothesis 4.2.1. We now introduce hypotheses directly on the $q_{ij}^{(k,m,N)}$ as the primary variables. Observe in particular that they will be assumed to be non-negative.

Hypothesis 4.2.19. *Assume that there exists $N_0 \in \mathbb{N}$ such that*

$$\sup_{N \geq N_0} \sum_{i,j,k=0,1} \sum_{m \geq 2} q_{ij}^{(k,m,N)} < \infty$$

for non-negative $q_{ij}^{(k,m,N)}$, $i, j, k = 0, 1$ and $m \geq 2$. We can use this condition as in Remark 4.2.2 to show that the rewritten rates can be used to determine a $\{0, 1\}^{\mathbb{Z}/N}$-valued Markov process ξ_t^N for $N \geq N_0$.

Hypothesis 4.2.20. *In the special case with no fixed kernel, i.e. where*

$$q_{00}^{(k,m,N)} = q_{10}^{(k,m,N)} \text{ and } q_{01}^{(k,m,N)} = q_{11}^{(k,m,N)} \iff q_{0j}^{(k,m,N)} = q_{1j}^{(k,m,N)}, j = 0, 1$$

(see Remark 4.2.3 and Remark 4.2.17) we assume that

$$\theta^{(N)} \overset{N\to\infty}{\to} \theta,$$
$$q_{0j}^{(k,m,N)} \overset{N\to\infty}{\to} q_{0j}^{(k,m)} \text{ for all } j,k = 0,1 \text{ and } m \geq 2$$

and

$$\lim_{N\to\infty} \sum_{j,k=0,1} \sum_{m\geq 2} q_{0j}^{(k,m,N)} = \sum_{j,k=0,1} \sum_{m\geq 2} q_{0j}^{(k,m)}. \tag{4.17}$$

Remark 4.2.21. *Observe that if we assume that the $q_{0j}^{(k,m,N)}, j,k = 0,1, m \geq 2$ were obtained from $\alpha_j^{(m,N)}, j = 0,1, m \geq 1$ as described earlier in Remark 4.2.17 and Remark 4.2.18, then (4.11) implies (4.17). Indeed, use for instance [12], Proposition 11.18 together with Remark 4.2.17.*

Notation 4.2.22. For $k = 0, 1$ and $a \in \mathbb{R}$ we let

$$F_k(a) = \begin{cases} 1 - a, & k = 0, \\ a, & k = 1. \end{cases}$$

By the above it remains to prove the following theorem. The claim of Theorem 4.2.9 will then follow immediately and Theorem 4.2.14 will follow using Corollary 4.2.24 below.

Theorem 4.2.23. *Suppose that $A(\xi_0^N) \to u_0$ in $\mathcal{C}_1$. Let the transition rates of $\xi^N(x)$ be as in (4.16) and $q_{ij}^{(k,m,N)}$ satisfying Hypothesis 4.2.19. Then the $\left(A(\xi_t^N) : t \geq 0\right)$ are C-tight as cadlag $\mathcal{C}_1$-valued processes and the $\left(\nu_t^N : t \geq 0\right)$ are C-tight as cadlag Radon measure valued processes with the vague topology. If $\left(A\left(\xi_t^{N_k}\right), \nu_t^{N_k}\right)_{t\geq 0}$ converges to $(u_t, \nu_t)_{t\geq 0}$, then $\nu_t(dx) = u_t(x)dx$ for all $t \geq 0$.*

For the special case with no fixed kernel we further obtain that if Hypothesis 4.2.20 holds, then the limit points of $A(\xi_t^N)$ are continuous $\mathcal{C}_1$-valued processes u_t which solve

$$\frac{\partial u}{\partial t} = \frac{\Delta u}{6} + \sum_{k=0,1} (1-2k) \sum_{m\geq 2, j=0,1} q_{0j}^{(k,m)} F_j(u) \left(F_{1-k}(u)\right)^{m-1} F_k(u) + \sqrt{2u(1-u)}\dot{W} \tag{4.18}$$

with initial condition u_0. If we assume additionally $< u_0, 1 >< \infty$, then u_t is the unique (in law) $[0,1]$-valued solution to the above SPDE.

In the next Corollary we assume there is no fixed kernel and the $q_{0j}^{(k,m)}, j,k = 0,1, m \geq 2$ are defined from the $\alpha_j^{(m)}, j = 0,1, m \geq 1$ as in Remark 4.2.17 and Remark 4.2.18 without the N's.

Corollary 4.2.24. *Under the assumption above, the SPDE (4.18) may be rewritten as*

$$u_t = \frac{\Delta u}{6} + (1-u)u\sum_{m=0}^{\infty}\alpha_0^{(m+1)}u^m - u(1-u)\sum_{m=0}^{\infty}\alpha_1^{(m+1)}(1-u)^m + \sqrt{2u(1-u)}\dot{W}.$$

Proof. Indeed first use the definition of $F_k(a)$ and collect terms appropriately. Then recall how we rewrote the transition rates in Corollary 4.2.16 and Remark 4.2.18 to obtain (4.16) from (4.14). Now analogously rewrite (4.12) as (4.18). □

Before we start proving the above we need some notation. In what follows we shall consider

$$e_\lambda(x) = \exp(\lambda|x|)$$

for $\lambda \in \mathbb{R}$ and we let

$$\mathcal{C} = \{f : \mathbb{R} \to [0,\infty) \text{ cont. with } |f(x)e_\lambda(x)| \to 0 \text{ as } |x| \to \infty \text{ for all } \lambda < 0\}$$

be the set of non-negative continuous functions with slower than exponential growth. Define

$$\| f \|_\lambda = \sup_x |f(x)e_\lambda(x)|$$

and give $\mathcal{C}$ the topology generated by the norms $(\|\cdot\|_\lambda \colon \lambda < 0)$.

Remark 4.2.25. *We work on the space $\mathcal{C}$ instead of $\mathcal{C}_1$ because in Section 4.4 we shall introduce functions $0 \le \psi_t^z(x) \le CN^{1/2}$ and shall show in Lemma 4.5.2(b) that they converge in $\mathcal{C}$ to the Brownian transition density $p\left(\frac{t}{3}, z-x\right)$. Finally, in Section 4.5 we shall derive estimates on p^{th}-moment differences of $\hat{A}(\xi_t)(z) \equiv A(\xi_t)(z) - <\xi_0, \psi_t^z>$, where $A(\xi_0) \to u_0$ in $\mathcal{C}$ to finally establish the tightness claim for the sequence of approximate densities $A(\xi^N)(x)$.*

Notation 4.2.26. For $x \in N^{-1}\mathbb{Z}$, $f : N^{-1}\mathbb{Z} \to \mathbb{R}$ and $\delta > 0$ we shall use

$$D(f,\delta)(x) = \sup\{|f(y) - f(x)| : |y - x| \le \delta, y \in N^{-1}\mathbb{Z}\}, \tag{4.19}$$

$$\Delta(f)(x) = \frac{N - \theta^{(N)}}{2c(N)N^{1/2}}\sum_{y\sim x}(f(y) - f(x)),$$

where we suppress the dependence on N.

4.3 An Approximate Martingale Problem

We shall now derive a graphical construction and evolution in time of our approximating processes ξ_t^N. The graphical construction uses independent families of i.i.d. Poisson processes:

$$\big(P_t(x;y) : x, y \in N^{-1}\mathbb{Z}\big) \text{ i.i.d. P.p. of rate } \frac{N-\theta^{(N)}}{2c(N)N^{1/2}}, \tag{4.20}$$

and for $m \geq 2, i, j, k = 0, 1$,

$$\Big(Q_t^{m,i,j,k}(x; y_1, \ldots, y_m; z) : x, y_1, \ldots, y_m, z \in N^{-1}\mathbb{Z}\Big)$$
$$\text{i.i.d. P.p. of rate } \frac{q_{ij}^{(k,m,N)}}{(2c(N))^m N^{m/2}} p(N(x-z)).$$

Note that we suppress the dependence on N in the family of Poisson processes $P_t(x;y)$ and $Q_t^{m,i,j,k}(x; y_1, \ldots, y_m; z)$.

At a jump of $P_t(x;y)$ the voter at x adopts the opinion of the voter at y provided that y has the opposite opinion.

At a jump of $Q_t^{m,i,j,k}(x; y_1, \ldots, y_m; z)$ the voter at x adopts the opinion $1-k$ provided that y_1 has opinion j, all of $y_2 \ldots, y_m$ have opinion $1-k$ and z has opinion i.

This yields the following SDE to describe the evolution in time of our approximating processes ξ_t^N:

$$\begin{aligned}
\xi_t^N(x) =& \xi_0^N(x) + \sum_{y\sim x} \int_0^t \left\{\delta_0\big(\xi_{s-}^N(x)\big)\,\delta_1\big(\xi_{s-}^N(y)\big) - \delta_1\big(\xi_{s-}^N(x)\big)\,\delta_0\big(\xi_{s-}^N(y)\big)\right\} \\
& \times dP_s(x;y) \\
& + \sum_{k=0,1}(1-2k) \sum_{m\geq 2, i,j=0,1} \sum_{y_1,\ldots,y_m\sim x} \sum_z \int_0^t \delta_k\big(\xi_{s-}^N(x)\big) \\
& \times \delta_j\big(\xi_{s-}^N(y_1)\big) \prod_{l=2}^m \delta_{1-k}\big(\xi_{s-}^N(y_l)\big)\,\delta_i\big(\xi_{s-}^N(z)\big)\,dQ_s^{m,i,j,k}(x; y_1, \ldots, y_m; z)
\end{aligned} \tag{4.21}$$

for all $x \in N^{-1}\mathbb{Z}$.

We now explain why the above system (4.21) has a unique solution. The problem with (4.21) is that although there is a first flip time for $\xi_t^N(x)$ (the jump rate there is bounded as can be shown using Hypothesis 4.2.19 and as the sum of Poisson processes is a Poisson process again, we have at most one flip at a time), what to do at this time depends on the states of the finite number of "communicating sites" $y, z \in \mathbb{Z}/N, y \sim x, |z-x| \leq C_p/N$. The equations for each of these sites will in turn involve its communicating sites. If we now try to go backwards in time to determine the configuration at $x \in \mathbb{Z}/N$, starting with $t > 0$, we may encounter an accumulation of jumps before time zero.

To see that the described problem cannot occur, we shall use the uniform boundedness of the flip-rates together with the finite interaction range $R \equiv \lceil N^{-1}(\sqrt{N} \vee C_p) \rceil$, where $\lceil x \rceil$ denotes the next largest integer.

Remark 4.3.1. *We have avoided random walk kernels p with infinite range for the fixed kernel interactions to simplify the analysis of the above jump equations (4.21).*

We show that up to time t, the evolution of $\xi_\cdot^N$ can be divided up into finite random "islands" that do not communicate with each other. Indeed, two regions of $\mathbb{Z}/N$ do not interact with each other up to time t, if we can find an intermediate region of length $2R$ where no flips occur in $[0, t]$. We can now partition $\mathbb{Z}/N$ into such regions. The sums of flips for each region up to time t are independent and can be bounded by i.i.d. Poisson random variables as follows. The flips in the region centered at $Z2R, Z \in \mathbb{Z}$ can be bounded by

$$P_t^Z \equiv \sum_{\substack{x \in \mathbb{Z}/N, \\ Z2R-R<x\leq Z2R+R}} P_{t,x},$$

where

$$P_{t,x} \equiv \sum_{y \sim x} P_t(x;y) + \sum_{m \geq 2, i,j,k=0,1} \sum_{y_1,\ldots,y_m \sim x, |z-x| \leq C_p/N} Q_t^{m,i,j,k}(x; y_1, \ldots, y_m; z).$$

Using (4.20) we obtain that each P_t^Z has mean

$$\begin{aligned} &t2RN \Bigg(2c(N)N^{1/2} \frac{N - \theta^{(N)}}{2c(N)N^{1/2}} \\ &\quad + \sum_{m\geq 2,i,j,k=0,1} \left(2c(N)N^{1/2}\right)^m \frac{q_{ij}^{(k,m,N)}}{(2c(N))^m N^{m/2}} \Bigg) \sum_{|z-x|\leq C_p/N} p(N(x-z)) \\ &= t2RN \left(N - \theta^{(N)} + \sum_{m\geq 2,i,j,k=0,1} q_{ij}^{(k,m,N)} \right). \end{aligned}$$

Thus by Hypothesis 4.2.19, $(P_t^Z)_{Z\in\mathbb{Z}}$ is a sequence of i.i.d. Poisson random variables of finite mean. Let X_Z be a sequence of random variables with $X_Z = 1$ if $P_t^Z > 0$ and $X_Z = 0$ if $P_t^Z = 0$. Then $(X_Z)_{Z\in\mathbb{Z}}$ is an i.i.d. sequence of Bernoulli variables with $\mathbb{P}(X_0 = 0) > 0$. In particular we can show that with probability one there exists a random sequence $\ldots, Z_{-2}, Z_{-1}, Z_0, Z_1, Z_2, \ldots$ such that $1 \leq |Z_i - Z_{i-1}| < \infty$ for all $i \in \mathbb{Z}$ and $X_{Z_i} = 0$. Hence, up until time t, we can partition

$$\mathbb{Z}/N = \cup_{i\in\mathbb{Z}} (Z_i 2R, Z_{i+1} 2R] \cap \mathbb{Z}/N$$

into finite regions that do not communicate with each other up to time t. For all x of each region we can now uniquely solve (4.21) on $[0, t]$. As the region has finite length, we only need to consider a finite number of sites. To see this, note that as $P_{t,x}$ is a Poisson random variable of finite mean for each x in the region, we can have at most a finite number of flips in each region up until time t. Now iterate on successive intervals of length t to uniquely solve the entire system for all times. The interested reader is referred to the proof of Proposition 2.1(a) in [4] for how to solve such systems.

Remark 4.3.2. *For an alternative proof of the partition in non-communicating islands the reader is referred to Theorem 2.1 in Durrett [6]. The ideas of [6] can be applied to our setup but as we only consider dimension* $d = 1$ *the more straightforward calculation given above was possible.*

Having solved the equation (4.21) it remains to ensure that the solution is the spin-flip system with rates $c(x, \xi^N)$ given by (4.16).

Recall the end of Remark 4.2.2 and the end of Hypothesis 4.2.19. By Theorem I.5.2 in [7] the process ξ^N constructed from the given rates is the unique in law solution to the martingale problem for Ω, where

$$\Omega f\big(\xi^N\big) = \sum_x c\big(x, \xi^N\big) \left(f(\xi^N_x) - f(\xi^N)\right)$$

with f in the space of finite cylinder functions on $\{0,1\}^{\mathbb{Z}/N}$.

Hence it remains to show that for the Markov process ξ^N constructed in (4.21),

$$\begin{aligned} & f(\xi^N_t) - f(\xi^N_0) - \int_0^t \Omega f(\xi^N_s) ds \\ & = f(\xi^N_t) - f(\xi^N_0) - \int_0^t \sum_x c(x, \xi^N_s) \left(f((\xi^N_s)_x) - f(\xi^N_s)\right) ds \end{aligned}$$

is a martingale for all f in the space of finite cylinder functions on $\{0,1\}^{\mathbb{Z}/N}$. Since f depends on finitely many coordinates, this is an exercise in stochastic calculus for jump processes, see Remark 4.3.4 below.

In what follows we shall often drop the superscripts w.r.t. N to simplify notation.

We now derive the approximate martingale problem. We take a test function $\phi : [0,\infty) \times N^{-1}\mathbb{Z} \to \mathbb{R}$ with $t \mapsto \phi_t(x)$ continuously differentiable and satisfying

$$\int_0^T < |\phi_s| + \phi_s^2 + |\partial_s \phi_s|, 1 > ds < \infty \tag{4.22}$$

(this condition ensures that the following integration and summation are well-defined). We apply integration by parts to $\xi_t(x)\phi_t(x)$, sum over x and multiply by $\frac{1}{N}$, to obtain for $t \leq T$ (recall that by Remark 4.2.6 $< \xi_t, \phi >=< \nu_t, \phi >$)

$$
\begin{aligned}
&< \nu_t, \phi_t > && (4.23)\\
&=< \nu_0, \phi_0 > + \int_0^t < \nu_s, \partial_s \phi_s > ds \\
&\quad + \frac{1}{N} \sum_x \sum_{y \sim x} \int_0^t \xi_{s-}(y) \left(\phi_s(x) - \phi_s(y) \right) dP_s(x;y) && (4.24)\\
&\quad + \frac{1}{N} \sum_x \sum_{y \sim x} \int_0^t \xi_{s-}(x) \phi_s(x) \left(dP_s(y;x) - dP_s(x;y) \right) && (4.25)\\
&\quad + \sum_{k=0,1} (1-2k) \sum_{m \geq 2, i,j=0,1} \frac{1}{N} \sum_x \sum_{y_1,\ldots,y_m \sim x} \sum_z \int_0^t \delta_k(\xi_{s-}(x)) \\
&\qquad \times \delta_j(\xi_{s-}(y_1)) \prod_{l=2}^m \delta_{1-k}(\xi_{s-}(y_l)) \, \delta_i(\xi_{s-}(z)) \, \phi_s(x) dQ_s^{m,i,j,k}(x; y_1, \ldots, y_m; z). \\
& && (4.26)
\end{aligned}
$$

The main ideas for analyzing terms (4.24) and (4.25) will become clear once we analyze term (4.26) in detail. The latter is the only term where calculations changed seriously compared to [9]. Hence, we shall only summarize the results for terms (4.24) and (4.25) in what follows.

We break term (4.24) into two parts, an average term and a fluctuation term and after proceeding as for term (3.1) in [9] we obtain

$$
(4.24) = \int_0^t < \nu_{s-}, \Delta(\phi_s) > ds + E_t^{(1)}(\phi),
$$

where

$$
E_t^{(1)}(\phi) \equiv \frac{1}{N} \sum_x \sum_{y \sim x} \int_0^t \xi_{s-}(y) \left(\phi_s(x) - \phi_s(y) \right) \left(dP_s(x;y) - d\langle P(x;y) \rangle_s \right).
$$

We have suppressed the dependence on N in $E_t^{(1)}(\phi)$. $E_t^{(1)}(\phi)$ is a martingale (recall that if $N \sim \text{Pois}(\lambda)$, then $N_t - \lambda t$ is a martingale with quadratic variation $\langle N \rangle_t = \lambda t$) with predictable brackets process given by

$$
d\big\langle E^{(1)}(\phi) \big\rangle_t \leq \left\| D\left(\phi_t, \frac{1}{\sqrt{N}} \right) \right\|_\lambda^2 < 1, e_{-2\lambda} > dt. \tag{4.27}
$$

Alternatively we also obtain the bound

$$
d\big\langle E^{(1)}(\phi) \big\rangle_t \leq 4 \parallel \phi_t \parallel_0 < |\phi_t| \, , 1 > dt \tag{4.28}
$$

with $\parallel \phi_t \parallel_0 = \sup_x |\phi_t(x)|$.

The second term (4.25) is a martingale which we shall denote by $M_t^{(N)}(\phi)$ (in what follows we shall drop the superscripts w.r.t. N and write $M_t(\phi)$). It can be analyzed as the martingale $Z_t(\phi)$ of (3.3) in [9]. We obtain in particular that

$$\langle M(\phi)\rangle_t = 2\frac{N-\theta^{(N)}}{N}\left\{\int_0^t < \xi_{s-}, \phi_s^2 > ds - \int_0^t < A(\xi_{s-}\phi_s), \xi_{s-}\phi_s > ds\right\}. \tag{4.29}$$

Using that

$$\begin{aligned}|A(\xi_{s-}\phi_s)(x)| &\equiv \left|\frac{1}{2c(N)N^{1/2}}\sum_{y\sim x}\xi_{s-}(y)\phi_s(y)\right| \\ &\le \frac{1}{2c(N)N^{1/2}}\sum_{y\sim x}|\phi_s(y)| \le \sup_{y\sim x}|\phi_s(y)|.\end{aligned}$$

we can further dominate $\langle M(\phi)\rangle_t$ by

$$\langle M(\phi)\rangle_t \le C(\lambda)\int_0^t \left(\|\phi_s\|_\lambda^2 < 1, e_{-2\lambda} >\right) \wedge \left(\|\phi_s\|_0 < \xi_{s-}, |\phi_s| >\right) ds. \tag{4.30}$$

We break the third term (4.26) into two parts, an average term and a fluctuation term. Recall Notation 4.2.22 and observe that if we only consider $a \in \{0,1\}$ we have $F_k(a) = \delta_k(a)$.

$$\begin{aligned}(4.26) = &\sum_{k=0,1}(1-2k)\sum_{m\ge 2,i,j=0,1}\frac{1}{N}\sum_x\sum_{y_1,\ldots,y_m\sim x}\sum_z\int_0^t \delta_k(\xi_{s-}(x))\,\delta_j(\xi_{s-}(y_1)) \\ &\times\prod_{l=2}^m \delta_{1-k}(\xi_{s-}(y_l))\,\delta_i(\xi_{s-}(z))\,\phi_s(x)\frac{q_{ij}^{(k,m,N)}}{(2c(N))^m N^{m/2}}p(N(x-z))ds \\ &+ E_t^{(3)}(\phi) \\ = &\sum_{k=0,1}(1-2k)\sum_{m\ge 2,i,j=0,1} q_{ij}^{(k,m,N)}\int_0^t \frac{1}{N}\sum_x\left(\frac{1}{2c(N)N^{1/2}}\sum_{y_1\sim x}\delta_j(\xi_{s-}(y_1))\right) \\ &\times\prod_{l=2}^m\left(\frac{1}{2c(N)N^{1/2}}\sum_{y_l\sim x}\delta_{1-k}(\xi_{s-}(y_l))\right)\left(\sum_z p(N(x-z))\delta_i(\xi_{s-}(z))\right) \\ &\times\delta_k(\xi_{s-}(x))\,\phi_s(x)ds + E_t^{(3)}(\phi) \\ = &\sum_{k=0,1}(1-2k)\sum_{m\ge 2,i,j=0,1} q_{ij}^{(k,m,N)}\int_0^t\frac{1}{N}\sum_x F_j(A(\xi_{s-})(x)) \\ &\times\left(F_{1-k}(A(\xi_{s-})(x))\right)^{m-1}F_i((p^N*\xi_{s-})(x))\delta_k(\xi_{s-}(x))\,\phi_s(x)ds + E_t^{(3)}(\phi) \\ = &\sum_{k=0,1}(1-2k)\sum_{m\ge 2,i,j=0,1} q_{ij}^{(k,m,N)}\int_0^t < (F_j\circ A(\xi_{s-})) \\ &\times\left(F_{1-k}\circ A(\xi_{s-})\right)^{m-1}\left(F_i\circ(p^N*\xi_{s-})\right)(\delta_k\circ\xi_{s-}), \phi_s > ds + E_t^{(3)}(\phi),\end{aligned}$$

where for $x \in \mathbb{Z}/N$ we set

$$\left(p^N * f\right)(x) \equiv \sum_{z\in\mathbb{Z}/N} p(N(x-z))f(z) \tag{4.31}$$

and

$$\begin{aligned} E_t^{(3)}(\phi) \equiv \sum_{k=0,1}(1-2k) &\sum_{m\geq 2,i,j=0,1} \frac{1}{N}\sum_x \sum_{y_1,\ldots,y_m\sim x}\sum_z \int_0^t \delta_k(\xi_{s-}(x)) \\ &\times \delta_j(\xi_{s-}(y_1))\prod_{l=2}^m \delta_{1-k}(\xi_{s-}(y_l))\,\delta_i(\xi_{s-}(z))\,\phi_s(x) \\ &\times \left(dQ_s^{m,i,j,k}(x;y_1,\ldots,y_m;z) - \frac{q_{ij}^{(k,m,N)}}{(2c(N))^m N^{m/2}} p(N(x-z))ds\right). \end{aligned}$$

We have suppressed the dependence on N in $E_t^{(3)}(\phi)$. Here, $E_t^{(3)}(\phi)$ is a martingale with predictable brackets process given by

$$\begin{aligned} \left\langle E^{(3)}(\phi)\right\rangle_t \leq &\sum_{m\geq 2,i,j,k=0,1} q_{ij}^{(k,m,N)}\frac{1}{N^2}\sum_x\prod_{l=0}^m\left(\sum_{y_l\sim x}\frac{1}{2c(N)N^{1/2}}\right) \\ &\times\left(\sum_z p(N(x-z))\right)\int_0^t \phi_s^2(x)ds \\ \leq &\sum_{m\geq 2,i,j,k=0,1} q_{ij}^{(k,m,N)}\frac{1}{N}\int_0^t < \phi_s^2, 1 > ds \\ \leq &\frac{1}{N}\sum_{m\geq 2,i,j,k=0,1} q_{ij}^{(k,m,N)}\int_0^t \parallel\phi_s\parallel_\lambda^2 < e_{-2\lambda}, 1 > ds. \end{aligned} \tag{4.32}$$

Taking the above together we obtain the following approximate semimartingale decomposition from (4.23).

$$\begin{aligned} < \nu_t, \phi_t > = < \nu_0, \phi_0 > + &\int_0^t < \nu_s, \partial_s\phi_s > ds \\ + &\int_0^t < \nu_{s-}, \Delta(\phi_s) > ds + E_t^{(1)}(\phi) + M_t(\phi) \\ + &\sum_{k=0,1}(1-2k)\sum_{m\geq 2,i,j=0,1} q_{ij}^{(k,m,N)}\int_0^t < (F_j\circ A(\xi_{s-})) \\ &\times (F_{1-k}\circ A(\xi_{s-}))^{m-1}\left(F_i\circ(p^N * \xi_{s-})\right)(\delta_k\circ\xi_{s-}), \phi_s > ds \\ + &E_t^{(3)}(\phi). \end{aligned} \tag{4.33}$$

Remark 4.3.3. *Note that this approximate semimartingale decomposition provides the link between our approximate densities and the limiting SPDE in (4.18) for the case with no fixed kernel. Indeed, uniqueness of the limit u_t of $A(\xi_t^N)$ will be derived by proving that u_t solves the martingale problem associated with the SPDE (4.18).*

Remark 4.3.4. *For all f in the space of finite cylinder functions on $\{0,1\}^{\mathbb{Z}/N}$ the Markov process $\xi(=\xi^N)$ constructed in (4.21) yields a martingale*

$$f(\xi_t)-f(\xi_0)-\int_0^t \Omega f(\xi_s)ds = f(\xi_t)-f(\xi_0)-\int_0^t \sum_x c(x,\xi_s)\left(f((\xi_s)_x)-f(\xi_s)\right)ds. \tag{4.34}$$

Indeed, every finite cylinder function on $\{0,1\}^{\mathbb{Z}/N}$, $f(\xi)=f(\xi(x_1),\ldots,\xi(x_n))$, $n\in\mathbb{N},\xi(x_i)\in\{0,1\},x_i\in\mathbb{Z}/N$ can be rewritten as a linear combination of functions of the form $g(\xi)\equiv\xi(x_{i_1})\cdots\xi(x_{i_m})$, where $m\in\mathbb{N},m\le n$ and $\{i_1,\ldots,i_m\}\subset\{1,\ldots,n\}$. By linearity we only need to consider functions of this form.

Now rewrite (4.21) as

$$\xi_t(x)=\xi_0(x)+\int_0^t c(x,\xi_{s-})\left(1-2\xi_{s-}(x)\right)ds + mart. \tag{4.35}$$

by breaking both integrals in (4.21) into an average term and a fluctuation term. Observe here that we can rewrite the sum of both average terms as in (4.35) by using for example

$$\delta_0(\xi_{s-}(x))\sum_{y\sim x}\delta_1(\xi_{s-}(y))\frac{N-\theta^{(N)}}{2c(N)N^{1/2}}=\delta_0(\xi_{s-}(x))\left(N-\theta^{(N)}\right)f_1(x,\xi_{s-})$$

and

$$\sum_z \delta_i(\xi_{s-}(z))\,p(N(x-z))=g_i(x,\xi_{s-})\,.$$

Now use the representation of the rates $c(x,\xi)$ from (4.16). Both fluctuation terms turn out to be martingales that only depend on the Poisson processes $P_t(x;y)$ and $Q_t^{m,i,j,k}(x;y_1,\ldots,y_m;z)$. Hence, for $x\neq x'$ the martingale terms are orthogonal.

For $m=2$, that is for $g(\xi)=\xi(x)\xi(x')$ with $x\neq x'$, we can now use the integration by parts formula (cf. Theorem VI.(38.3) in Rogers and Williams [13]), the orthogonality of the martingale terms of $\xi_t(x)$ and $\xi_t(x')$ and

$$g(\xi_x)-g(\xi)=(1-\xi(x))\xi(x')-\xi(x)\xi(x')=(1-2\xi(x))\xi(x')$$

to obtain that (4.34) is a martingale for $f=g$. Now iterate the above reasoning to obtain the claim for all $m\in\mathbb{N}$.

4.4 Green's Function Representation

Analogous to [9], define a test function

$$\psi_t^z(x) \geq 0 \text{ for } t \geq 0, x, z \in N^{-1}\mathbb{Z}$$

as the unique solution, satisfying (4.22) and s.t.

$$\frac{\partial}{\partial t}\psi_t^z = \Delta\psi_t^z,$$
$$\psi_0^z(x) = \frac{N^{1/2}}{2c(N)}1(x \sim z)$$

with

$$\Delta\psi_t^z(x) = \frac{N-\theta^{(N)}}{2c(N)N^{1/2}}\sum_{y\sim x}(\psi_t^z(y) - \psi_t^z(x)) \tag{4.36}$$

as in (4.19). Note that ψ_0^z was chosen s.t. $< \nu_t, \psi_0^z >= A(\xi_t)(z)$ and that we suppress the dependence on N.

Next observe that Δ is the generator of a simple random walk $X_t \in N^{-1}\mathbb{Z}$, jumping at rate $\frac{N-\theta^{(N)}}{2c(N)N^{1/2}}\left(2c(N)N^{1/2}\right) = N - \theta^{(N)} = (1+o(1))N$ with symmetric steps of variance $\frac{1}{N}\left(\frac{1}{3}+o(1)\right)$, where we used that $c(N) \overset{N\to\infty}{\to} 1$. Here $o(1)$ denotes some function that satisfies $o(1) \to 0$ for $N \to \infty$. Define

$$\bar{\psi}_t^z(x) = N\mathbb{P}(X_t = x|X_0 = z)$$

then

$$< \psi_0^z, \bar{\psi}_t^x > = \frac{1}{N}\sum_y \psi_0^z(y)\bar{\psi}_t^x(y) = \frac{1}{N}\sum_y \frac{N^{1/2}}{2c(N)}1(y\sim z)N\mathbb{P}(X_t = y|X_0 = x) \tag{4.37}$$
$$= \frac{N^{1/2}}{2c(N)}\sum_{y\sim z}\mathbb{P}(X_t = y|X_0 = x) = \mathbb{E}_x[\psi_0^z(X_t)] = \psi_t^z(x).$$

As we shall see later in Lemma 4.5.2(b), when linearly interpolated, the functions $\psi_t^z(x)$ and $\bar{\psi}_t^z(x)$ converge to $p\left(\frac{t}{3}, z-x\right)$ (the proof follows), where

$$p_t(x) = \frac{1}{\sqrt{2\pi t}}e^{-\frac{x^2}{2t}} \text{ is the Brownian transition density.}$$

The next Lemma gives some information on the test functions ψ and $\bar{\psi}$ from above. Later on, this will provide us with estimates necessary for establishing tightness.

Lemma 4.4.1. *There exists $N_0 < \infty$ s.t. for $N \geq N_0, T \geq 0, z \in N^{-1}\mathbb{Z}, \lambda \geq 0$,*

(a) $<\psi_t^z, 1> = <\bar{\psi}_t^z, 1> = 1$ *and* $\| \psi_t^z \|_0 \leq CN^{1/2}$ *for all* $t \geq 0$.

(b) $< e_\lambda, \psi_t^z + \bar{\psi}_t^z > \leq C(\lambda, T) e_\lambda(z)$ *for all* $t \leq T$,

(c) $\| \psi_t^z \|_\lambda \leq C(\lambda, T) \left(N^{1/2} \wedge t^{-2/3} \right) e_\lambda(z)$ *for all* $t \leq T$,

(d) $< \left| \bar{\psi}_t^z - \bar{\psi}_s^z \right|, 1 > \leq 2N|t-s|$ *for all* $s, t \geq 0$.

If we further restrict ourselves to $N \geq N_0$, $N^{-3/4} \leq s < t \leq T, y, z \in N^{-1}\mathbb{Z}$, $|y-z| \leq 1$, we get

(e)
$$\| \psi_t^z - \psi_t^y \|_\lambda \leq C(\lambda, T) e_\lambda(z) \left(|z-y|^{1/2} t^{-1} + N^{-1/2} t^{-3/2} \right),$$

(f)
$$\| \psi_t^z - \psi_s^z \|_\lambda \leq C(\lambda, T) e_\lambda(z) \left(|t-s|^{1/2} s^{-3/2} + N^{-1/2} s^{-3/2} \right),$$

(g)
$$\left| \left| D\left(\psi_t^z, N^{-1/2} \right)(\cdot) \right| \right|_\lambda \leq C(\lambda, T) e_\lambda(z) N^{-1/4} t^{-1}.$$

Proof. First we shall derive an explicit description for the test functions ψ_t^z and $\bar{\psi}_t^z$. We proceed as at the beginning of Section 4 in [9] by using that Δ as in (4.36) is the generator of a simple random walk.

Let $(Y_i)_{i=1,2,\ldots}$ be i.i.d. and uniformly distributed on $(jN^{-1} : 0 < |j| \leq \sqrt{N})$. Set

$$\rho(t) = \mathbb{E}\left[e^{itY_1}\right] \text{ and } S_k = \sum_{i=1}^{k} Y_i. \tag{4.38}$$

Note that $E[Y_1^2] = \frac{1+o(1)}{3N}$, where $o(1) \to 0$ for $N \to \infty$. Similarly, $E[Y_1^4] = \frac{1+o(1)}{5N^2}$, where $o(1)$ may change from line to line.

The relation between the test functions $\psi_t^z, \bar{\psi}_t^z$ and S_k is as follows.

$$\psi_t^z(x) = \mathbb{E}_x[\psi_0^z(X_t)] = \sum_{k=0}^{\infty} \frac{((N-\theta^{(N)})t)^k}{k!} e^{-((N-\theta^{(N)})t)} N\mathbb{P}(S_{k+1} = x - z), \tag{4.39}$$

$$\bar{\psi}_t^z(x) = N\mathbb{P}(X_t = x | X_0 = z) = \sum_{k=0}^{\infty} \frac{((N-\theta^{(N)})t)^k}{k!} e^{-((N-\theta^{(N)})t)} N\mathbb{P}(S_k = x - z).$$

Now we can start proving the above Lemma.

(a) follows as in the proof of Lemma 3(a), [9], using that $\mathbb{P}(S_k = x) \leq CN^{-1/2}$ for all $x \in N^{-1}\mathbb{Z}, k \geq 1$.

(b) follows as in the proof of Lemma 3(b), [9], where we shall use the bound

$$\mathbb{E}\left[e^{\mu|Y_1|}\right] \leq \exp\left\{5\mu^2 \frac{1}{N}\right\}$$

for all $\mu \geq 0$ to obtain the claim.

(c) Following the proof of Lemma 3(c) in [9], we first show that we have, for $k \in \mathbb{N}$ and $|x| \geq 1$, $\mathbb{P}(S_k = x) \leq \frac{1}{N}\mathbb{P}(S_k \geq |x| - 1)$, which we can use to obtain $\mathbb{P}(S_k = x) \leq \frac{1}{N}e^{-\mu(|x|-1)}\exp\{5k\mu^2\frac{1}{N}\}$.

Substituting this bound into (4.39) gives for any $\mu \geq 0$

$$\psi_t^z(x) \leq C(\mu, T)\exp\{-\mu|x - z|\} \tag{4.40}$$

for all $t \leq T$ and $|x - z| \geq 1$.

From (4.39) we further have for N big enough

$$\psi_t^z(x) \equiv \mathbb{E}\left[p\left(\frac{(1 + o(1))(\mathcal{P}_t + 1)}{3N}, x - z\right)\right] + E(N, t, x - z),$$

where $\mathcal{P}_t \sim \text{Pois}((N - \theta^{(N)})t)$. Using Corollary B.0.2 we get as in the proof of [9], Lemma 3(c),

$$|E(N, t, x)| \leq C\frac{1}{N}\left(1 + t^{-3/2}\right) \qquad \text{for } N^{-3/4} \leq t.$$

Here we used that for $\mathcal{P} \sim \text{Pois}(r), r > 0$ we have

$$\mathbb{E}[(\mathcal{P} + 1)^a] \leq C(a)r^a \text{ for all } a < 0.$$

(This is obviously true for $0 < r < 1$. For $r \geq 1$ fixed, prove the claim first for all $a \in \mathbb{Z}$. Then extend this result to general $a < 0$ by an application of Hölder's inequality.)

Using the trivial bound $p(t, x) \leq Ct^{-1/2}$ we get from the above

$$\psi_t^z(x) \leq C(T)t^{-2/3} \qquad \text{for } N^{-3/4} \leq t \leq T.$$

Finally, we obtain

$$\begin{aligned}
\|\,\psi_t^z\,\|_\lambda &= \sup_x\left\{|\psi_t^z(x)|\, e^{\lambda|x|}\right\} \\
&\overset{(4.40)}{\leq} \sup_{\{x:|x-z|\geq 1\}}\left\{C(\lambda, T)e^{-\lambda|x-z|}e^{\lambda|x|}\right\} \\
&\qquad \vee \sup_{\{x:|x-z|<1, N^{-3/4}\leq t\leq T\}}\left\{C(T)t^{-2/3}e^{\lambda|x|}\right\} \\
&\qquad \vee \sup_{\{x:|x-z|<1, 0\leq t\leq N^{-3/4}\}}\left\{|\psi_t^z(x)|\, e^{\lambda|x|}\right\} \\
&\overset{(a)}{\leq} \left\{C(\lambda, T)e^{\lambda|z|}\right\} \vee \left\{1\left(N^{-3/4} \leq t \leq T\right) C(T)t^{-2/3}e^1 e^{\lambda|z|}\right\} \\
&\qquad \vee \left\{1\left(0 \leq t \leq N^{-3/4}\right) CN^{1/2}e^1 e^{\lambda|z|}\right\} \\
&\leq C(\lambda, T)\left(N^{1/2} \wedge t^{-2/3}\right) e_\lambda(z) \qquad \text{for all } t \leq T.
\end{aligned}$$

This proves part (c).

(d) follows along the lines of the proof of [9], Lemma 3(d).

(e) For the remaining parts (e)-(g) we fix $N^{-3/4} \le s < t \le T, y, z \in N^{-1}\mathbb{Z}$, $|y-z| \le 1$. For part (e) we follow the reasoning of the proof of [9], Lemma 3(e). The only change occurs in the derivation of the last estimate. In summary, we find as in [9] that

$$\| \psi_t^z - \psi_t^y \|_0 \le C(T) \left(|z-y| t^{-1} + N^{-1} t^{-3/2} \right). \tag{4.41}$$

Now recall (4.40) with $\mu = 2\lambda$ to get $\psi_t^z(x) + \psi_t^y(x) \le C(\lambda, T) \exp\{-2\lambda |x-z|\}$ for $|x-z| \ge 1, |x-y| \ge 1, |y-z| \le 1$ and thus in particular for $|x-z| \ge 2, |y-z| \le 1$. This yields

$$\begin{aligned}
\| \psi_t^z - \psi_t^y \|_\lambda \le & \sup_{\{x:|x-z|<2\}} \| \psi_t^z - \psi_t^y \|_0 \, e_\lambda(x) \\
& + \sup_{\{x:|x-z|\ge 2\}} \left\{ C(\lambda, T) \, \| \psi_t^z - \psi_t^y \|_0^{1/2} \, e^{-\lambda|x-z|} e_\lambda(x) \right\} \\
\le & C(\lambda, T) e_\lambda(z) \left(\| \psi_t^z - \psi_t^y \|_0 + \| \psi_t^z - \psi_t^y \|_0^{1/2} \right) \\
\le & C(\lambda, T) e_\lambda(z) \left(\left(|z-y| t^{-1} + N^{-1} t^{-3/2} \right) \right. \\
& \left. + \left(|z-y| t^{-1} + N^{-1} t^{-3/2} \right)^{1/2} \right) \\
\le & C(\lambda, T) e_\lambda(z) \left(|z-y|^{1/2} t^{-1} + N^{-1/2} t^{-3/2} \right).
\end{aligned}$$

This proves (e).

(f) The proof of part (f) follows analogously to the proof of part (e), with changes as suggested in the proof of [9], Lemma 3(f).

(g) Finally, to prove part (g), use part (e), $\psi_t^z(y) = \psi_t^y(z)$ (see (4.39)) and the definition of

$$D\left(\psi_t^z, N^{-1/2}\right)(x) = \sup\left\{ |\psi_t^z(y) - \psi_t^z(x)| : |x-y| \le N^{-1/2}, y \in N^{-1}\mathbb{Z} \right\}$$

to get

$$\begin{aligned}
& \| D\left(\psi_t^z, N^{-1/2}\right)(\cdot) \|_\lambda \\
\le & \sup_{\{x:|x-z|<2\}} \left\{ \sup_{y:|x-y|\le N^{-1/2}} \{ |\psi_t^z(y) - \psi_t^z(x)| \} \, e^{\lambda|x|} \right\} \\
& + \sup_{\{x:|x-z|\ge 2\}} \left\{ \sup_{y:|x-y|\le N^{-1/2}} \{ |\psi_t^z(y) - \psi_t^z(x)| \} \, e^{\lambda|x|} \right\} \\
\overset{(4.40)}{\le} & \, C(\lambda) \sup_{\{x:|x-z|<2\}} \left\{ \sup_{y:|x-y|\le N^{-1/2}} \{ |\psi_t^z(y) - \psi_t^z(x)| \} \, e^{\lambda|z|} \right\} \\
& + C(\lambda, T) \sup_{\{x:|x-z|\ge 2\}} \left\{ \sup_{y:|x-y|\le N^{-1/2}} \left\{ |\psi_t^z(y) - \psi_t^z(x)|^{1/2} \right\} e^{-\lambda|x-z|} e^{\lambda|x|} \right\}.
\end{aligned}$$

Next use that $\psi_t^a(b) = \psi_t^b(a)$ to get as a further upper bound

$$\begin{aligned}
&C(\lambda) \sup_{\{x:|x-z|<2\}} \left\{ \sup_{y:|x-y|\le N^{-1/2}} \{|\psi_t^y(z) - \psi_t^x(z)|\} e^{\lambda|z|} \right\} \\
&\quad + C(\lambda,T) \sup_{\{x:|x-z|\ge 2\}} \left\{ \sup_{y:|x-y|\le N^{-1/2}} \left\{|\psi_t^y(z) - \psi_t^x(z)|^{1/2}\right\} e^{\lambda|z|} \right\} \\
&\overset{(4.41)}{\le} C(\lambda,T)e_\lambda(z) \sup_{x,y:|x-y|\le N^{-1/2}} \Big\{ \Big(|x-y|t^{-1} + N^{-1}t^{-3/2}\Big) \\
&\quad + \Big(|x-y|t^{-1} + N^{-1}t^{-3/2}\Big)^{1/2} \Big\} \\
&\le C(\lambda,T)e_\lambda(z)N^{-1/4}t^{-1},
\end{aligned}$$

where we used $N^{-3/4} < t \le T$. This finishes the proof of (g) and it also finishes the proof of the Lemma. □

The following Corollary uses the results of Lemma 4.4.1 to obtain estimates that we shall need later.

Corollary 4.4.2. *There exists $N_0 < \infty$ s.t. for $N \ge N_0$, $0 \le \delta \le u \le t \le T$ and $y, z \in N^{-1}\mathbb{Z}, \lambda \ge 0$, we have*

(a) $\int_u^t \| \psi_{t-s}^z \|_\lambda \, ds \le C(\lambda,T)(t-u)^{1/3} e_\lambda(z)$
and $\int_0^t \| \psi_{t-s}^z \|_\lambda^2 \, ds \le C(\lambda,T) N^{1/4} e_{2\lambda}(z)$.

(b) For $|y-z| \le 1$ and $\delta \le t - N^{-3/4}$ we further have
$\sup_{0\le s\le\delta} \| \psi_{t-s}^z - \psi_{t-s}^y \|_\lambda$
$\le C(\lambda,T)e_\lambda(z)\left\{|z-y|^{1/2}(t-\delta)^{-1} + N^{-1/2}(t-\delta)^{-3/2}\right\}$.

(c) We also have $\int_\delta^t \| \psi_{t-s}^z - \psi_{t-s}^y \|_\lambda \, ds \le C(\lambda,T)\left(e_\lambda(z) + e_\lambda(y)\right)(t-\delta)^{1/3}$.

(d) For $N^{-3/4} \le u - \delta$ we have
$\sup_{0\le s\le\delta} \| \psi_{t-s}^z - \psi_{u-s}^z \|_\lambda$
$\le C(\lambda,T)e_\lambda(z)\left\{(t-u)^{1/2}(u-\delta)^{-3/2} + N^{-1/2}(u-\delta)^{-3/2}\right\}$.

(e) Finally, we have $\int_\delta^u \| \psi_{t-s}^z - \psi_{u-s}^z \|_\lambda \, ds \le C(\lambda,T)e_\lambda(z)(u-\delta)^{1/3}$.

Proof. The proof is a combination of the results of Lemma 4.4.1.

(a) We have for $n = 1, 2$ and $0 \le u \le t$ by Lemma 4.4.1(c)

$$\int_u^t \left|\left|\psi_{t-s}^z\right|\right|_\lambda^n ds \le C(\lambda,T) \int_u^t N^{n/2} \wedge (t-s)^{-2n/3} ds \; e_{n\lambda}(z).$$

For $n = 1$ further bound the integrand by $(t-s)^{-2/3}$, for $n = 2$ and $u = 0$ use the above integrand to obtain the claim.

(b) follows from Lemma 4.4.1(e).

(c) We further have by Lemma 4.4.1(c)

$$\int_\delta^t \| \psi^z_{t-s} - \psi^y_{t-s} \|_\lambda \, ds \leq C(\lambda, T) \left(e_\lambda(z) + e_\lambda(y) \right) \int_\delta^t (t-s)^{-2/3} ds.$$

(d) follows from Lemma 4.4.1(f).
(e) Using Lemma 4.4.1(c) once more, we get

$$\int_\delta^u \| \psi^z_{t-s} - \psi^z_{u-s} \|_\lambda \, ds \leq C(\lambda, T) e_\lambda(z) \int_\delta^u (t-s)^{-2/3} + (u-s)^{-2/3} ds,$$

which concludes the proof after some basic calculations. □

We shall need the following technical Lemma.

Lemma 4.4.3. *For $f : N^{-1}\mathbb{Z} \to [0, \infty)$ with $< f, 1 >< \infty, \lambda \in \mathbb{R}$ we have*

(a) $< \nu_s, \psi^z_{t-s} >=< A(\xi_s), \bar{\psi}^z_{t-s} >$,

(b) $|< \nu_t, f > - < A(\xi_t), f >| \leq C(\lambda) \, \| D(f, N^{-1/2}) \|_\lambda$.

Proof. *(a)* follows easily from

$$\begin{aligned} < \nu_s, \psi^z_{t-s} > =< \xi_s, \psi^z_{t-s} > &= \frac{1}{N} \sum_x \xi_s(x) \psi^z_{t-s}(x) = \frac{1}{N} \sum_x \xi_s(x) \psi^x_{t-s}(z) \\ &\stackrel{(4.37)}{=} \frac{1}{N} \sum_x \xi_s(x) < \psi^x_0, \bar{\psi}^z_{t-s} > \\ &= \frac{1}{N} \sum_x \xi_s(x) \frac{1}{N} \sum_y \frac{N^{1/2}}{2c(N)} 1(y \sim x) \bar{\psi}^z_{t-s}(y) \\ &= \frac{1}{N} \sum_y \left\{ \sum_x \frac{1}{2c(N)N^{1/2}} 1(y \sim x) \xi_s(x) \right\} \bar{\psi}^z_{t-s}(y) \\ &= \frac{1}{N} \sum_y A(\xi_s)(y) \bar{\psi}^z_{t-s}(y) =< A(\xi_s), \bar{\psi}^z_{t-s} > . \end{aligned}$$

Part (b) follows as in the proof of Lemma 5(b) in [9]. Observe in particular that $< \nu_t, e_{-\lambda} >\leq C(\lambda)$ as will be shown before and in (4.44) below.
Taken all together this finishes the proof. □

Next use the test function

$$\phi_s \equiv \psi^x_{t-s} \text{ for } s \leq t$$

in the semimartingale decomposition (4.33) and observe that ϕ satisfies (4.22). Here the initial condition is chosen so that $< \nu_t, \phi_t >=< \nu_t, \psi^x_0 >= A(\xi_t)(x)$.

The test function chosen in [9] at the beginning of page 526, namely $\phi_s = e^{\theta_c(t-s)} \psi^x_{t-s}$ was chosen so that the drift term $< \nu_s, \theta_c \phi_s > ds$ of the semimartingale decomposition (2.9) in [9] would cancel out with the drift term

$< \nu_s, \partial_s \phi_s > ds$. As we have multiple coefficients, this is not possible. Also, it turned out that the calculations become easier once we consider time differences in Section 4.5 to follow.

With the above choice we obtain, for a fixed value of t, an approximate Green's function representation for $A(\xi_t)$, namely

$$\begin{aligned} A(\xi_t)(x) = &< \nu_0, \psi_t^x > + \int_0^t < \nu_s, \left(-\Delta\psi_{t-s}^x\right) > ds \\ &+ \int_0^t < \nu_{s-}, \Delta\left(\psi_{t-s}^x\right) > ds + E_t^{(1)}\left(\psi_{t-\cdot}^x\right) + M_t\left(\psi_{t-\cdot}^x\right) \\ &+ \sum_{k=0,1}(1-2k) \sum_{m\geq 2,i,j=0,1} q_{ij}^{(k,m,N)} \int_0^t < (F_j \circ A(\xi_{s-})) \\ &\times (F_{1-k}\circ A(\xi_{s-}))^{m-1}\left(F_i \circ (p^N * \xi_{s-})\right)(\delta_k \circ \xi_{s-}), \psi_{t-s}^x > ds \\ &+ E_t^{(3)}(\psi_{t-\cdot}^x). \end{aligned} \tag{4.42}$$

The following Lemma is stated analogously to Lemma 4 of [9]. Parts (a) and (c) will follow easily in our setup and so the only significant statement will be part (b).

Lemma 4.4.4. *Suppose that the initial conditions satisfy $A(\xi_0) \to u_0$ in $\mathcal{C}$ as $N \to \infty$. Then for $T \geq 0, p \geq 2, \lambda > 0$,*

(a) $\mathbb{E}\left[\sup_{t\leq T} < \nu_t, e_{-\lambda} >^p\right] \leq C(\lambda, p)$.

(b) We further have

$$\mathbb{E}\left[\left|E_t^{(1)}\left(\psi_{t-\cdot}^z\right)\right|^p \vee \left|E_t^{(3)}\left(\psi_{t-\cdot}^z\right)\right|^p\right] \leq C(\lambda, p, T)\left(1 + C_Q^{p/2}\right) N^{-p/16} e_{\lambda p}(z)$$

for all $t \leq T$ and N big enough, where we set

$$C_Q \equiv \sup_{N\geq N_0} \sum_{m\geq 2,i,j,k=0,1} q_{ij}^{(k,m,N)}. \tag{4.43}$$

(c) Finally, $\| \mathbb{E}[A(\xi_t)] \|_{-\lambda p} \leq 1$ for all $t \leq T$.

Proof. First observe that we have $\xi_t \in \{0,1\}^{\mathbb{Z}/N}$ and $0 \leq A(\xi_t) \leq 1$. Therefore, parts (a) and (c) follow immediately. Indeed, for (a) we only need to observe that

$$\begin{aligned} 0 \leq < \nu_t, e_{-\lambda} > = < \xi_t, e^{-\lambda|\cdot|} > = \frac{1}{N}\sum_x \xi_t(x) e^{-\lambda|x|} \leq \frac{1}{N} \sum_{x:x\in\mathbb{Z}/N} e^{-\lambda|x|} \\ \leq \frac{2}{N}\sum_{j=0}^{\infty} e^{-\lambda j/N} = \frac{2}{N}\frac{1}{1-e^{-\lambda/N}} \overset{N\to\infty}{\to} \frac{2}{\lambda}. \end{aligned}$$

Note in particular that we showed that

$$< e_{-\lambda}, 1 > \leq \frac{C}{\lambda} \text{ for all } \lambda > 0,\ N = N(\lambda) \text{ big enough,} \tag{4.44}$$

which will prove useful later.

For (c) we further have

$$\| \mathbb{E}[A(\xi_t)] \|_{-\lambda p} = \sup_x |\mathbb{E}[A(\xi_t)(x)]| \, e^{-\lambda p|x|} \leq \sup_x e^{-\lambda p|x|} \leq 1.$$

It only remains to show that (b) holds.

(b) First observe that $C_Q < \infty$ by Hypothesis 4.2.19.

We shall apply a Burkholder-Davis-Gundy inequality in the form

$$\mathbb{E}\left[\sup_{s\leq t} |X_s|^p\right] \leq C(p)\mathbb{E}\left[\langle X\rangle_t^{p/2} + \sup_{s\leq t} |X_s - X_{s-}|^p\right] \tag{4.45}$$

for a cadlag martingale X with $X_0 = 0$ (this inequality may be derived from its discrete time version, see Burkholder [1], Theorem 21.1).

To get an upper bound on the second term of the r.h.s. of (4.45) for the martingales we consider, observe that the largest possible jumps of the martingales $E_t^{(1)}(\psi_{t-\cdot}^z)$ respectively $E_t^{(3)}(\psi_{t-\cdot}^z)$ are bounded a.s. by $CN^{-1/2}$. Indeed,

$$E_t^{(1)}(\psi_{t-\cdot}^z) = \frac{1}{N}\sum_x \sum_{y\sim x} \int_0^t \xi_{s-}(y) \left(\psi_{t-s}^z(x) - \psi_{t-s}^z(y)\right) (dP_s(x;y) - d\langle P(x;y)\rangle_s)$$

and thus, using Lemma 4.4.1(a), the maximal jump size is bounded by

$$\frac{1}{N} 2 \sup_{t\leq T} \| \psi_t^z \|_0 \leq \frac{C}{N^{1/2}} \tag{4.46}$$

(the maximal number of jumps at a fixed time is 1). The bound on the maximal jump size of $E_t^{(3)}(\psi_{t-\cdot}^z)$ follows analogously.

Now choose $t \leq T$. We shall start with $E_t^{(3)}(\psi_{t-\cdot}^z)$. By (4.45), (4.46) and (4.32) we have

$$\begin{aligned}
&\mathbb{E}\left[\left|E_t^{(3)}(\psi_{t-\cdot}^z)\right|^p\right] \\
&\leq C(p)\left(\sum_{m\geq 2, i,j,k=0,1} q_{ij}^{(k,m,N)} \frac{1}{N}\int_0^t \left|\left|\psi_{t-s}^z\right|\right|_\lambda^2 < e_{-2\lambda}, 1 > ds\right)^{p/2} + C(p)N^{-p/2} \\
&\overset{(4.44)}{\leq} C(\lambda,p)C_Q^{p/2} N^{-p/2}\left\{\left(\int_0^t \left|\left|\psi_{t-s}^z\right|\right|_\lambda^2 ds\right)^{p/2} + 1\right\}.
\end{aligned}$$

By Corollary 4.4.2(a) this is bounded from above by

$$C(\lambda,p,T)C_Q^{p/2}N^{-p/2}\left\{N^{p/8} e_{\lambda p}(z) + 1\right\} = C(\lambda,p,T)C_Q^{p/2}N^{-3p/8} e_{\lambda p}(z).$$

It remains to investigate $E_t^{(1)}(\psi^z_{t-\cdot})$. Here (4.45), (4.46), (4.27) and (4.28) yield

$$\begin{aligned}
&\mathbb{E}\left[\left|E_t^{(1)}(\psi^z_{t-\cdot})\right|^p\right] \\
&\leq C(p)\left(\int_0^t \left[\| \psi^z_{t-s}\|_0 < \psi^z_{t-s}, 1 >\right] \wedge \left[\left\|D\left(\psi^z_{t-s}, \frac{1}{\sqrt{N}}\right)\right\|_\lambda^2 < 1, e_{-2\lambda} >\right] ds\right)^{p/2} \\
&\quad + C(p)N^{-p/2}.
\end{aligned}$$

This in turn is bounded from above by

$$C(p)\left(\int_0^t \left[C(T)(t-s)^{-2/3}\right] \wedge \left[\left\|D\left(\psi^z_{t-s}, \frac{1}{\sqrt{N}}\right)\right\|_\lambda^2 C(\lambda)\right] ds\right)^{p/2} + C(p)N^{-p/2},$$

where we used Lemma 4.4.1(a), (c) and (4.44). To apply Lemma 4.4.1(g) to the second part of the integrand, we need to ensure that $N^{-3/4} \leq t-s$. As $N^{-3/4} \leq N^{-3/8}$ we get as a further upper bound

$$\begin{aligned}
&C(p)\left(\int_0^{N^{-3/8}} C(T)s^{-2/3}ds + \int_{N^{-3/8}\wedge t}^t \left(C(\lambda,T)e_\lambda(z)N^{-1/4}s^{-1}\right)^2 C(\lambda)ds\right)^{p/2} \\
&+ C(p)N^{-p/2} \\
&\leq C(\lambda,p,T)e_{\lambda p}(z)\left\{\left(\left(N^{-3/8}\right)^{1/3} + N^{-1/2}\left(N^{-3/8}\right)^{-1}\right)^{p/2} + N^{-p/2}\right\} \\
&\leq C(\lambda,p,T)N^{-p/16}e_{\lambda p}(z).
\end{aligned}$$

This finishes the proof. □

4.5 Tightness

In what follows we shall derive estimates on p^{th}-moment differences of

$$\hat{A}(\xi_t)(z) \equiv A(\xi_t)(z) - < \nu_0, \psi^z_t > .$$

Recall the assumption

$$A(\xi_0) \to u_0 \text{ in } \mathcal{C}$$

from Theorem 4.2.9 resp. Theorem 4.2.23 from the beginning. Also note that Lemma 4.5.2(b) to come will yield that $\psi^z_t(x)$ converges to $p\left(\frac{t}{3}, z-x\right)$. The estimates of Lemma 4.5.1 and the convergence of ψ^z_t taken together will be sufficient to show C-tightness of the approximate densities $A(\xi_t)(z)$ at the end of this Section.

Lemma 4.5.1. *For $0 \leq s \leq t \leq T, y, z \in N^{-1}\mathbb{Z}, |t-s| \leq 1, |y-z| \leq 1, \lambda > 0$ and $p \geq 2$ we have*

$$\begin{aligned}&\mathbb{E}\Big[\Big|\hat{A}(\xi_t)(z) - \hat{A}(\xi_s)(y)\Big|^p\Big]\\&\leq C(\lambda, p, T)\left(1 + C_Q^p\right) e_{\lambda p}(z)\left(|t-s|^{p/24} + |z-y|^{p/24} + N^{-p/24}\right).\end{aligned}$$

Proof. Fix $s, t, T, y, z, \lambda, p$ as in the statement. We decompose the increment $\hat{A}(\xi_t)(z) - \hat{A}(\xi_s)(y)$ into a space increment $\hat{A}(\xi_t)(z) - \hat{A}(\xi_t)(y)$ and a time increment $\hat{A}(\xi_t)(y) - \hat{A}(\xi_s)(y)$.

We consider first the space differences. From the Green's function representation (4.42), the estimates obtained in Lemma 4.4.4(b) for the error terms $E^{(1)}$ and $E^{(3)}$ and the linearity of $M_t(\phi)$ and $E_t^{(1)}(\phi), E_t^{(3)}(\phi)$ in ϕ, we get

$$\begin{aligned}&\mathbb{E}\Big[\Big|\hat{A}(\xi_t)(z) - \hat{A}(\xi_t)(y)\Big|^p\Big]\\&\leq C(\lambda, p, T)\left(1 + C_Q^{p/2}\right) N^{-p/16} e_{\lambda p}(z) + \mathbb{E}\big[\big|M_t\big(\psi_{t-\cdot}^z - \psi_{t-\cdot}^y\big)\big|^p\big]\\&\quad + \mathbb{E}\Big[\Big|\sum_{k=0,1}(1-2k)\sum_{m\geq 2, i,j=0,1} q_{ij}^{(k,m,N)}\int_0^t < (F_j \circ A(\xi_{s-}))\\&\qquad \times (F_{1-k}\circ A(\xi_{s-}))^{m-1}\left(F_i \circ (p^N * \xi_{s-})\right)(\delta_k \circ \xi_{s-}), \left(\psi_{t-s}^z - \psi_{t-s}^y\right) > ds\Big|^p\Big].\end{aligned}$$

Recall definition (4.31) and observe that $0 \leq \left(p^N * \xi_{s-}\right)(x) \leq 1$ follows from $\xi_{s-} \in \{0,1\}^{\mathbb{Z}/N}$. Use this and $0 \leq A(\xi_{s-})(x) \leq 1$ together with the definition of F_k from Notation 4.2.22 to get

$$\begin{aligned}&\mathbb{E}\Big[\Big|\hat{A}(\xi_t)(z) - \hat{A}(\xi_t)(y)\Big|^p\Big]\\&\leq C(\lambda, p, T)\left(1 + C_Q^{p/2}\right) N^{-p/16} e_{\lambda p}(z) + \mathbb{E}\big[\big|M_t\big(\psi_{t-\cdot}^z - \psi_{t-\cdot}^y\big)\big|^p\big]\\&\quad + \mathbb{E}\Big[\Big(\sum_{m\geq 2, i,j,k=0,1} q_{ij}^{(k,m,N)}\\&\qquad \times \int_0^t < (F_{1-k}\circ A(\xi_{s-}))(\delta_k \circ \xi_{s-}), \left|\psi_{t-s}^z - \psi_{t-s}^y\right| > ds\Big)^p\Big]\\&\leq C(\lambda, p, T)\left(1 + C_Q^{p/2}\right) N^{-p/16} e_{\lambda p}(z) + \mathbb{E}\big[\big|M_t\big(\psi_{t-\cdot}^z - \psi_{t-\cdot}^y\big)\big|^p\big]\\&\quad + C_Q^p \mathbb{E}\Big[\Big(\int_0^t < A(\xi_{s-}) + \xi_{s-}, \left|\psi_{t-s}^z - \psi_{t-s}^y\right| > ds\Big)^p\Big].\end{aligned} \tag{4.47}$$

Note that this is the main step to see why the fixed kernel interaction does not impact our results on tightness.

In what follows, we shall employ a similar strategy to the proof of Lemma 6 in [9] to obtain estimates on the above. We nevertheless give full calculations as we proceeded in a different logical order to highlight the ideas for obtaining bounds. Minor changes in the exponents of our bounds ensued, both due to the different logical order and the different setup.

Let us first derive a bound on $\mathbb{E}\big[\big|M_t\big(\psi^z_{t-\cdot}-\psi^y_{t-\cdot}\big)\big|^p\big]$. Using the Burkholder-Davis-Gundy inequality (4.45) from above and observing that the jumps of the martingales $M_t(\psi^x_{t-\cdot})$ are bounded a.s. by $CN^{-1/2}$ we have for any $0\le\delta\le t$

$$\begin{aligned}
&\mathbb{E}\big[\big|M_t\big(\psi^z_{t-\cdot}-\psi^y_{t-\cdot}\big)\big|^p\big] \qquad (4.48)\\
&\overset{(4.30)}{\le} C(\lambda,p)\mathbb{E}\Bigg[\Bigg(\int_0^\delta \|\psi^z_{t-s}-\psi^y_{t-s}\|^2_\lambda < 1, e_{-2\lambda}> ds\\
&\qquad + \int_\delta^t \|\psi^z_{t-s}-\psi^y_{t-s}\|_0 < \xi_{s-}, \big|\psi^z_{t-s}-\psi^y_{t-s}\big| > ds\Bigg)^{p/2}\Bigg] + C(p)N^{-p/2}\\
&\overset{(4.44)}{\le} C(\lambda,p)\mathbb{E}\Bigg[\Bigg(T\sup_{0\le s\le\delta}\|\psi^z_{t-s}-\psi^y_{t-s}\|^2_\lambda \frac{1}{\lambda}\\
&\qquad + \int_\delta^t \|\psi^z_{t-s}-\psi^y_{t-s}\|_0 < \xi_{s-}, \big|\psi^z_{t-s}-\psi^y_{t-s}\big| > ds\Bigg)^{p/2}\Bigg] + C(p)N^{-p/2}.
\end{aligned}$$

Now observe that by Lemma 4.4.3(a) and Lemma 4.4.1(a),

$$\begin{aligned}
< \xi_{s-}, \big|\psi^z_{t-s}-\psi^y_{t-s}\big| > &\le < \xi_{s-}, \psi^z_{t-s}+\psi^y_{t-s} > \qquad (4.49)\\
&= < A(\xi_{s-}), \bar\psi^z_{t-s}+\bar\psi^y_{t-s} >\\
&\le < 1, \bar\psi^z_{t-s}+\bar\psi^y_{t-s} >\\
&= 2.
\end{aligned}$$

We can therefore apply the estimates from Corollary 4.4.2(b) to the first term in (4.48) and Corollary 4.4.2(c) to the second term, assuming $\delta\le\big(t-N^{-3/4}\big)\vee 0$ and using $|y-z|\le 1$ to obtain

$$\begin{aligned}
&\mathbb{E}\big[\big|M_t\big(\psi^z_{t-\cdot}-\psi^y_{t-\cdot}\big)\big|^p\big]\\
&\le C(\lambda,p,T)e_{\lambda p}(z)\Big\{|z-y|^{p/2}(t-\delta)^{-p} + N^{-p/2}(t-\delta)^{-3p/2} + (t-\delta)^{p/6}\Big\}\\
&\quad + C(p)N^{-p/2}.
\end{aligned}$$

Now set

$$\delta = t - \Big(\Big(|z-y|^{1/4}\vee N^{-1/4}\Big)\wedge t\Big)$$

and observe that $\delta\le\big(t-N^{-3/4}\big)\vee 0$ follows. We obtain $t-\delta=\big(|z-y|^{1/4}\vee N^{-1/4}\big)\wedge t$ and

$$\begin{aligned}
|z-y|^{1/4}\le N^{-1/4} \Rightarrow\ & |z-y|^{p/2}(t-\delta)^{-p} + N^{-p/2}(t-\delta)^{-3p/2} + (t-\delta)^{p/6} \qquad (4.50)\\
&\le |z-y|^{p/2}|z-y|^{-p/4} + N^{-p/2}N^{3p/8} + N^{-p/24}\\
&= |z-y|^{p/4} + N^{-p/8} + N^{-p/24},\\
|z-y|^{1/4} > N^{-1/4} \Rightarrow\ & |z-y|^{p/2}(t-\delta)^{-p} + N^{-p/2}(t-\delta)^{-3p/2} + (t-\delta)^{p/6}\\
&\le |z-y|^{p/2}|z-y|^{-p/4} + N^{-p/2}N^{3p/8} + |z-y|^{p/24}\\
&= |z-y|^{p/4} + N^{-p/8} + |z-y|^{p/24}.
\end{aligned}$$

Plugging this back in the above estimate we finally have

$$\mathbb{E}\big[\big|M_t\big(\psi^z_{t-\cdot} - \psi^y_{t-\cdot}\big)\big|^p\big] \le C(\lambda,p,T)e_{\lambda p}(z)\left\{|z-y|^{p/24} + N^{-p/24}\right\}.$$

Next we shall get a bound on the last term of (4.47). Recall that $<\xi_t,\phi>=<\nu_t,\phi>$. We get

$$\begin{aligned}
&\mathbb{E}\Big[\Big(\int_0^t < A(\xi_{s-}) + \xi_{s-}, \big|\psi^z_{t-s} - \psi^y_{t-s}\big| > ds\Big)^p\Big] \\
&\le C(p)\left\{\mathbb{E}\left[\left(\int_0^\delta < A(\xi_{s-}) + \nu_{s-}, e_{-\lambda} > ds \sup_{0\le s\le\delta} \| \psi^z_{t-s} - \psi^y_{t-s}\|_\lambda\right)^p\right]\right. \\
&\qquad \left. + \mathbb{E}\left[\left(\int_\delta^t < A(\xi_{s-}) + \nu_{s-}, e_{-\lambda} > \| \psi^z_{t-s} - \psi^y_{t-s}\|_\lambda \, ds\right)^p\right]\right\}.
\end{aligned}$$

Now use that

$$< A(\xi_{s-}) + \nu_{s-}, e_{-\lambda} > = < A(\xi_{s-}) + \xi_{s-}, e_{-\lambda} > \le < 2, e_{-\lambda} > \overset{(4.44)}{\le} C(\lambda) \quad (4.51)$$

to obtain that the above is bounded by

$$\begin{aligned}
&C(p)\left\{\left(TC(\lambda)\sup_{0\le s\le\delta} \| \psi^z_{t-s} - \psi^y_{t-s}\|_\lambda\right)^p + \left(\int_\delta^t C(\lambda)\, \| \psi^z_{t-s} - \psi^y_{t-s}\|_\lambda \, ds\right)^p\right\} \\
&\le C(\lambda,p,T)e_{\lambda p}(z)\left\{|z-y|^{p/2}(t-\delta)^{-p} + N^{-p/2}(t-\delta)^{-3p/2}\right. \\
&\qquad\qquad \left. +(t-\delta)^{p/3}\right\},
\end{aligned}$$

where we used Corollary 4.4.2(b),(c) and $|y-z|\le 1$. Here we assumed $\delta \le \big(t - N^{-3/4}\big)\vee 0$ when we applied Corollary 4.4.2(b). Now choose $\delta = t - \big(\big(|z-y|^{1/4}\vee N^{-1/4}\big)\wedge t\big) \le \big(t-N^{-3/4}\big)\vee 0$ as before. Reasoning as in (4.50), we get

$$C(\lambda,p,T)e_{\lambda p}(z)\left(N^{-p/8} + |z-y|^{p/12}\right)$$

as an upper bound.

Now we can take all the above bounds together and plug them back into (4.47) to obtain (recall that $|z-y|\le 1$)

$$\begin{aligned}
&\mathbb{E}\left[\left|\hat{A}(\xi_t)(z) - \hat{A}(\xi_t)(y)\right|^p\right] \\
&\le C(\lambda,p,T)\left(1 + C_Q^{p/2}\right)N^{-p/16}e_{\lambda p}(z) \\
&\quad + C(\lambda,p,T)\left(1 + C_Q^p\right)e_{\lambda p}(z)\left\{|z-y|^{p/24} + N^{-p/24}\right\} \\
&\le C(\lambda,p,T)\left(1 + C_Q^{p/2} + C_Q^p\right)e_{\lambda p}(z)\left(|z-y|^{p/24} + N^{-p/24}\right).
\end{aligned}$$

Next we derive a similar bound on the time differences. We start by subtracting the two Green's function representations again, this time for the time differences, using (4.42) and Lemma 4.4.4(b) for the error terms.

$$
\begin{aligned}
&\mathbb{E}\Big[\Big|\hat{A}(\xi_t)(z) - \hat{A}(\xi_u)(z)\Big|^p\Big] \qquad (4.52)\\
&\le C(\lambda,p,T)\left(1 + C_Q^{p/2}\right) N^{-p/16} e_{\lambda p}(z) + \mathbb{E}\big[\big|M_t\big(\psi^z_{t-\cdot}\big) - M_u\big(\psi^z_{u-\cdot}\big)\big|^p\big]\\
&\quad + \mathbb{E}\Big[\Big|\sum_{k=0,1}(1-2k)\sum_{m\ge 2, i,j=0,1} q_{ij}^{(k,m,N)}\\
&\qquad \times \Big\{\int_0^t < (F_j \circ A(\xi_{s-}))\,(F_{1-k}\circ A(\xi_{s-}))^{m-1}\,\big(F_i \circ (p^N * \xi_{s-})\big)\\
&\qquad\qquad \times (\delta_k \circ \xi_{s-})\,, \psi^z_{t-s} > ds\\
&\qquad\quad - \int_0^u < (F_j \circ A(\xi_{s-}))\,(F_{1-k}\circ A(\xi_{s-}))^{m-1}\,\big(F_i \circ (p^N * \xi_{s-})\big)\\
&\qquad\qquad \times (\delta_k \circ \xi_{s-})\,, \psi^z_{u-s} > ds\Big\}\Big|^p\Big]\\
&\le C(\lambda,p,T)\left(1 + C_Q^{p/2}\right) N^{-p/16} e_{\lambda p}(z) + \mathbb{E}\big[\big|M_t\big(\psi^z_{t-\cdot}\big) - M_u\big(\psi^z_{u-\cdot}\big)\big|^p\big]\\
&\quad + \mathbb{E}\Big[\Big(\sum_{m\ge 2, i,j,k=0,1} q_{ij}^{(k,m,N)}\\
&\qquad \times \Big\{\int_u^t < (F_j \circ A(\xi_{s-}))\,(F_{1-k}\circ A(\xi_{s-}))^{m-1}\,\big(F_i \circ (p^N * \xi_{s-})\big)\\
&\qquad\qquad \times (\delta_k \circ \xi_{s-})\,, \psi^z_{t-s} > ds\\
&\qquad\quad + \int_0^u < (F_j \circ A(\xi_{s-}))\,(F_{1-k}\circ A(\xi_{s-}))^{m-1}\,\big(F_i \circ (p^N * \xi_{s-})\big)\\
&\qquad\qquad \times (\delta_k \circ \xi_{s-})\,, \big|\psi^z_{t-s} - \psi^z_{u-s}\big| > ds\Big\}\Big)^p\Big]\\
&\le C(\lambda,p,T)\left(1 + C_Q^{p/2}\right) N^{-p/16} e_{\lambda p}(z) + \mathbb{E}\big[\big|M_t\big(\psi^z_{t-\cdot}\big) - M_u\big(\psi^z_{u-\cdot}\big)\big|^p\big]\\
&\quad + \mathbb{E}\Big[\Big(\sum_{m\ge 2, i,j,k=0,1} q_{ij}^{(k,m,N)} \Big\{\int_u^t < (F_{1-k}\circ A(\xi_{s-}))\,(\delta_k \circ \xi_{s-})\,, \psi^z_{t-s} > ds\\
&\qquad\quad + \int_0^u < (F_{1-k}\circ A(\xi_{s-}))\,(\delta_k \circ \xi_{s-})\,, \big|\psi^z_{t-s} - \psi^z_{u-s}\big| > ds\Big\}\Big)^p\Big]\\
&\le C(\lambda,p,T)\left(1 + C_Q^{p/2}\right) N^{-p/16} e_{\lambda p}(z) + \mathbb{E}\big[\big|M_t\big(\psi^z_{t-\cdot}\big) - M_u\big(\psi^z_{u-\cdot}\big)\big|^p\big]\\
&\quad + C_Q^p\,\mathbb{E}\Big[\Big(\int_u^t < A(\xi_{s-}) + \xi_{s-}, \psi^z_{t-s} > ds\\
&\qquad + \int_0^u < A(\xi_{s-}) + \xi_{s-}, \big|\psi^z_{t-s} - \psi^z_{u-s}\big| > ds\Big)^p\Big].
\end{aligned}
$$

For the martingale term we further get via the Burkholder-Davis-Gundy inequality (4.45)

$$
\begin{aligned}
&\mathbb{E}\big[\big|M_t\big(\psi^z_{t-\cdot}\big) - M_u\big(\psi^z_{t-\cdot}\big)\big|^p\big] \\
&\le C(p)\left\{\mathbb{E}\big[\big|M_t\big(\psi^z_{t-\cdot}\big) - M_u\big(\psi^z_{t-\cdot}\big)\big|^p\big] + \mathbb{E}\big[\big|M_u\big(\psi^z_{t-\cdot}\big) - M_u\big(\psi^z_{u-\cdot}\big)\big|^p\big]\right\} \\
&\le C(p)\mathbb{E}\Big[\big|\big\langle M_\cdot\big(\psi^z_{t-\cdot}\big)\big\rangle_t - \big\langle M_\cdot\big(\psi^z_{t-\cdot}\big)\big\rangle_u\big|^{p/2}\Big] \\
&\quad + C(p)\mathbb{E}\Big[\big|\big\langle M_\cdot\big(\psi^z_{t-\cdot} - \psi^z_{u-\cdot}\big)\big\rangle_u\big|^{p/2}\Big] + C(p)N^{-p/2} \\
&\le C(\lambda,p)\mathbb{E}\left[\left(\int_u^t \|\psi^z_{t-s}\|_0 < \xi_{s-}, \psi^z_{t-s} > ds\right)^{p/2}\right] \\
&\quad + C(\lambda,p)\left(\int_0^{\delta\wedge u} \|\psi^z_{t-s} - \psi^z_{u-s}\|^2_\lambda < 1, e_{-2\lambda} > ds\right)^{p/2} \\
&\quad + C(\lambda,p)\mathbb{E}\left[\left(\int_{\delta\wedge u}^u \|\psi^z_{t-s} - \psi^z_{u-s}\|_0 < \xi_{s-}, \big|\psi^z_{t-s} - \psi^z_{u-s}\big| > ds\right)^{p/2}\right] \\
&\quad + C(p)N^{-p/2},
\end{aligned}
$$

where we used equation (4.30) to bound the first and second term. Using (4.44) and reasoning as in (4.49) the above can further be bounded by

$$
\begin{aligned}
&\mathbb{E}\big[\big|M_t\big(\psi^z_{t-\cdot}\big) - M_u\big(\psi^z_{u-\cdot}\big)\big|^p\big] \\
&\le C(\lambda,p)\left(\int_u^t \|\psi^z_{t-s}\|_0\, ds\right)^{p/2} + C(\lambda,p,T)\sup_{0\le s\le\delta\wedge u}\|\psi^z_{t-s} - \psi^z_{u-s}\|^p_\lambda \\
&\qquad + C(\lambda,p)\left(\int_{\delta\wedge u}^u \|\psi^z_{t-s} - \psi^z_{u-s}\|_0\, ds\right)^{p/2} + C(p)N^{-p/2}.
\end{aligned}
$$

Under the assumption $N^{-3/4}\wedge u \le u - (\delta\wedge u)$ we obtain from Corollary 4.4.2(a), (d), (e) that

$$
\begin{aligned}
&\mathbb{E}\big[\big|M_t\big(\psi^z_{t-\cdot}\big) - M_u\big(\psi^z_{u-\cdot}\big)\big|^p\big] \qquad (4.53)\\
&\le C(\lambda,p,T)e_{\lambda p}(z)\Big\{(t-u)^{p/6} + \Big(|t-u|^{p/2} + N^{-p/2}\Big)\left(u-(\delta\wedge u)\right)^{-3p/2} \\
&\qquad + \left(u-(\delta\wedge u)\right)^{p/6} + N^{-p/2}\Big\}.
\end{aligned}
$$

Finally observe that with

$$
\delta = u - \left(\left(|t-u|^{1/4}\vee N^{-1/4}\right)\wedge u\right)
$$

we get $N^{-3/4}\wedge u \le u-\delta$ and by proceeding as in (4.50) we obtain

$$
\begin{aligned}
&\mathbb{E}\big[\big|M_t\big(\psi^z_{t-\cdot}\big) - M_u\big(\psi^z_{u-\cdot}\big)\big|^p\big] \\
&\le C(\lambda,p,T)e_{\lambda p}(z)\Big\{(t-u)^{p/6} + |t-u|^{p/24} + N^{-p/24} + N^{-p/2}\Big\}.
\end{aligned}
$$

Finally, we can bound the last expectation of the last line of (4.52) by using

$$< A(\xi_{t-s}) + \xi_{s-}, \psi^z_{t-s} > \leq < 1+1, \psi^z_{t-s} >= 2.$$

Here the last equality followed from Lemma 4.4.1(a). We thus obtain as an upper bound on the last expectation of the last line of (4.52),

$$C(p)\left\{|t-u|^p + \mathbb{E}\Bigg[\bigg(\int_0^u < A(\xi_{s-}) + \nu_{s-}, \left|\psi^z_{t-s} - \psi^z_{u-s}\right| > ds\bigg)^p\Bigg]\right\}.$$

We further have for the second term

$$\begin{aligned}
&\mathbb{E}\Big[\Big(\int_0^u < A(\xi_{s-}) + \nu_{s-}, \left|\psi^z_{t-s} - \psi^z_{u-s}\right| > ds\Big)^p\Big] \\
&\leq C(p)\left\{\mathbb{E}\left[\left(\int_0^{\delta\wedge u} < A(\xi_{s-}) + \nu_{s-}, e_{-\lambda} > ds \sup_{0\leq s\leq\delta} \| \psi^z_{t-s} - \psi^z_{u-s} \|_\lambda\right)^p\right]\right. \\
&\qquad \left. + \mathbb{E}\left[\left(\int_{\delta\wedge u}^u < A(\xi_{s-}) + \nu_{s-}, e_{-\lambda} > \| \psi^z_{t-s} - \psi^z_{u-s} \|_\lambda \, ds\right)^p\right]\right\} \\
&\stackrel{(4.51)}{\leq} C(\lambda,p,T)\left\{\left(\sup_{0\leq s\leq\delta\wedge u} \| \psi^z_{t-s} - \psi^z_{u-s} \|_\lambda\right)^p + \left(\int_{\delta\wedge u}^u \| \psi^z_{t-s} - \psi^z_{u-s} \|_\lambda \, ds\right)^p\right\} \\
&\leq C(\lambda,p,T) e_{\lambda p}(z)\left\{(t-u)^{p/2}(u-(\delta\wedge u))^{-3p/2} + N^{-p/2}(u-(\delta\wedge u))^{-3p/2}\right. \\
&\qquad\qquad \left. + (u-(\delta\wedge u))^{p/3}\right\},
\end{aligned}$$

where we assumed $N^{-3/4} \wedge u \leq u - (\delta \wedge u)$ when we applied Corollary 4.4.2(d) together with Corollary 4.4.2(e) in the last line. Now reason as from (4.53) on to obtain

$$C(\lambda,p,T) e_{\lambda p}(z)\left\{|t-u|^{p/24} + N^{-p/24}\right\}$$

as an upper bound.

Taking all bounds together we have for the time differences from (4.52)

$$\begin{aligned}
&\mathbb{E}\left[\left|\hat{A}(\xi_t)(z) - \hat{A}(\xi_u)(z)\right|^p\right] \\
&\leq C(\lambda,p,T)\left(1 + C_Q^{p/2}\right) N^{-p/16} e_{\lambda p}(z) \\
&\quad + C(\lambda,p,T) e_{\lambda p}(z)\left\{(t-u)^{p/6} + |t-u|^{p/24} + N^{-p/24} + N^{-p/2}\right\} \\
&\quad + C(p) C_Q^p\left\{|t-u|^p + C(\lambda,p,T) e_{\lambda p}(z)\left(|t-u|^{p/24} + N^{-p/24}\right)\right\} \\
&\leq C(\lambda,p,T)\left(1 + C_Q^{p/2} + C_Q^p\right) e_{\lambda p}(z)\left\{|t-u|^{p/24} + N^{-p/24}\right\}.
\end{aligned}$$

The bounds on the space difference and the time difference taken together complete the proof. □

We now show that these moment estimates imply C-tightness of the approximate densities. We shall start including dependence on N again to clarify the tightness argument. First define

$$\tilde{A}(\xi_t^N)(z) = \hat{A}(\xi_t^N)(z) \text{ on the grid } z \in N^{-1}\mathbb{Z}, t \in N^{-2}\mathbb{N}_0.$$

Linearly interpolate first in z and then in t to obtain a continuous $\mathcal{C}$-valued process. Note in particular that we can use Lemma 4.5.1 to show that for $0 \le s \le t \le T, |t-s| \le 1$ and $y, z \in \mathbb{R}, |y-z| \le 1$,

$$\begin{aligned}&\mathbb{E}\Big[\Big|\tilde{A}(\xi_t^N)(z) - \tilde{A}(\xi_s^N)(y)\Big|^p\Big] \\ &\le C(\lambda, p, T)\left(1 + C_Q^p\right) e_{\lambda p}(z)\left(|t-s|^{p/48} + |z-y|^{p/24}\right)\end{aligned}$$

for $\lambda > 0, p \ge 2$ arbitrarily fixed.

The next Lemma shows that $\tilde{A}(\xi_t^N)$ and $\hat{A}(\xi_t^N)$ remain close. The advantage of using $\tilde{A}(\xi_t^N)$ is that it is continuous.

Using Kolmogorov's continuity theorem (see for instance Corollary 1.2 in Walsh [15]) on compacts $R_1^{(i_1,i_2)} \equiv \{(t,x) \in \mathbb{R}^+ \times \mathbb{R} : (t,x) \in (i_1,i_2) + [0,1]^2\}$ for $i_1 \in \mathbb{N}_0, i_2 \in \mathbb{Z}$ we obtain tightness of $\tilde{A}(\xi_t^N)(x)$ in the space of continuous functions on $\left\{(t,x) : (t,x) \in R_1^{(i_1,i_2)}\right\}$. Indeed, we can use the Arzelà-Ascoli theorem. With arbitrary high probability, part (ii) of Corollary 1.2 of [15] provides a uniform (in N) modulus of continuity for all $N \ge N_0$. Pointwise boundedness follows from the boundedness of $A(\xi_t^N)(x)$ together with Lemma 4.5.2(b) below. Now use a diagonalization argument to obtain tightness of $(\tilde{A}(\xi_t^N)(x) : t \in \mathbb{R}^+, x \in \mathbb{R})_{N\in\mathbb{N}}$ in the space of continuous functions from $\mathbb{R}^+ \times \mathbb{R}$ to $\mathbb{R}$ equipped with the topology of uniform convergence on compact sets. Next observe that if we consider instead the space of continuous functions from $\mathbb{R}^+$ to the space of continuous functions from $\mathbb{R}$ to $\mathbb{R}$, both equipped with the topology of uniform convergence on compact sets, tightness of $(\tilde{A}(\xi_t^N)(x) : t \in \mathbb{R}^+, x \in \mathbb{R})_{N\in\mathbb{N}}$ in the former space is equivalent to tightness of $(\tilde{A}(\xi_t^N)(\cdot) : t \in \mathbb{R}^+)_{N\in\mathbb{N}}$ in the latter.

Finally, tightness of $(A(\xi_t^N) : t \in \mathbb{R}^+)_{N\in\mathbb{N}}$ as cadlag $\mathcal{C}_1$-valued processes (recall that $0 \le A(\xi_t^N)(x) \le 1$ by construction) and also the continuity of all weak limit points follow from the next Lemma.

Lemma 4.5.2. *For any* $\lambda > 0, T < \infty$ *we have*

(a) $\mathbb{P}\Big(\sup_{t\le T} \| \tilde{A}(\xi_t^N) - \hat{A}(\xi_t^N) \|_{-\lambda} \ge 7N^{-1/4}\Big) \to 0$ *as* $N \to \infty$.

(b) $\sup_{t\le T} \| < \nu_0^N, \psi_t^\cdot > -P_{t/3}u_0 \|_{-\lambda} \to 0$ *as* $N \to \infty$.

Proof. The proof is very similar to the proof of Lemma 7 in [9]. We shall only give some additional steps for part (a) to complement the proof of the given reference.

(a) For $0 \leq s \leq t$ we have

$$\begin{aligned}\| < \nu_0^N, \psi_t^{\cdot} > - < \nu_0^N, \psi_s^{\cdot} > \|_{-\lambda} &= \sup_z \left| < A(\xi_0^N), \bar{\psi}_t^z - \bar{\psi}_s^z > \right| e^{-\lambda|z|} \\ &\leq \sup_z < 1, \left| \bar{\psi}_t^z - \bar{\psi}_s^z \right| > \\ &\leq 2N|t-s|.\end{aligned}$$

Here we used Lemma 4.4.3(a) in the first line, $0 \leq A(\xi_0^N) \leq 1$ in the second line and Lemma 4.4.1(d) in the last. Hence, this only changes by $O(N^{-1})$ between the (time-)grid points in $N^{-2}\mathbb{N}_0$. We obtain that

$$\begin{aligned}&\mathbb{P}\Big(\sup_{t \leq T} \| \tilde{A}(\xi_t^N) - \hat{A}(\xi_t^N) \|_{-\lambda} \geq 7N^{-1/4}\Big) \\ &= \mathbb{P}\big(\exists t \in [0,T] \cap N^{-2}\mathbb{N}_0, s \in [0,T], |s-t| \leq N^{-2} \text{ s.t.} \\ &\qquad \| \hat{A}(\xi_t^N) - \hat{A}(\xi_s^N) \|_{-\lambda} \geq 7N^{-1/4}\big) \\ &\leq \mathbb{P}\big(\exists t \in [0,T] \cap N^{-2}\mathbb{N}_0, s \in [0,T], |s-t| \leq N^{-2} \text{ s.t.} \\ &\qquad \| A(\xi_t^N) - A(\xi_s^N) \|_{-\lambda} + \big\| < \nu_0^N, \psi_t^{\cdot} - \psi_s^{\cdot} > \big\|_{-\lambda} \geq 7N^{-1/4}\big) \\ &\leq \mathbb{P}\big(\exists t \in [0,T] \cap N^{-2}\mathbb{N}_0, s \in [0,T], |s-t| \leq N^{-2} \text{ s.t.} \\ &\qquad \| A(\xi_t^N) - A(\xi_s^N) \|_{-\lambda} \geq 6N^{-1/4}\big)\end{aligned}$$

for N big enough.

Next note that the value of $A(\xi_t^N)(x)$ changes only at jump times of $P_t(x;y)$ or $Q_t^{m,i,j.k}(x; y_1, \ldots, y_m; z), i,j,k = 0,1, m \geq 2$ for some $y \sim x$ respectively for some $y_1, \ldots, y_m \sim x$ and arbitrary $z \in N^{-1}\mathbb{Z}$ and that each jump of $A(\xi_t^N)$ is by definition of $A(\xi_t^N)$ bounded by $N^{-1/2}$. Then, writing $\mathcal{P}(a)$ for a Poisson variable with mean a, we get as a further bound on the above

$$\begin{aligned}&\sum_{l \in \mathbb{Z}} \mathbb{P}\big(\exists z \in N^{-1}\mathbb{Z} \cap (l, l+1], \exists t \in [0,T] \cap N^{-2}\mathbb{N}_0, \exists s \in [t, t+N^{-2}] \text{ with} \\ &\qquad \big\{ \big| A(\xi_t^N)(z) - A(\xi_s^N)(z) \big| \wedge \big| A(\xi_{t+N^{-2}}^N(z) - A(\xi_s^N)(z) \big| \big\} \geq N^{-1/4} e^{\lambda(|l|-1)}\big) \\ &\leq \sum_{l \in \mathbb{Z}} N(N^2 T) \mathbb{P}\Bigg(CN^{-1/2} \Bigg(\sum_{y \sim 0} P_{N^{-2}}(0;y) \\ &\qquad + \sum_{i,j,k=0,1,m \geq 2} \sum_{y_1,\ldots,y_m \sim 0} \sum_u Q_{N^{-2}}^{m,i,j,k}(0; y_1, \ldots, y_m; u)\Bigg) \geq N^{-1/4} e^{\lambda(|l|-1)}\Bigg) \\ &\leq \sum_{l \in \mathbb{Z}} C(T) N^3 \mathbb{P}\Big(CN^{-1/2} \mathcal{P}\Big(N^{-2}\Big(\Big(N - \theta^{(N)}\Big) + C_Q\Big)\Big) \geq N^{-1/4} e^{\lambda(|l|-1)}\Big) \\ &\leq \sum_{l \in \mathbb{Z}} C(T) N^3 \mathbb{P}\Big(\big(\mathcal{P}\big(N^{-2}(N + C_Q)\big)\big)^p \geq C N^{p/4} e^{\lambda p(|l|-1)}\Big)\end{aligned}$$

for some $p > 0$. Now apply Chebyshev's inequality. Choose $p > 0$ such that $3 - p/4 < 0$. Then the resulting sum is finite and goes to zero for $N \to \infty$.

(b) The proof of part (b) follows as the proof of Lemma 7(b) of [9]. □

4.6 Characterizing Limit Points

To conclude the proof of Theorem 4.2.9, Theorem 4.2.14 and Theorem 4.2.23 we can proceed as in Section 4 in [9], except for the proof of weak uniqueness of (4.12) respectively (4.18).

We shall give a short overview in what follows. The interested reader is referred to [9] for complete explanations.

In short, Lemma 4.4.3(b) implies for all $\phi \in \mathcal{C}_c$ that

$$\sup_t \left| < \nu_t^N, \phi > - < A(\xi_t^N), \phi > \right| \leq C(\lambda) \parallel D(\phi, N^{-1/2}) \parallel_\lambda \overset{N \to \infty}{\to} 0. \qquad (4.54)$$

The C-tightness of $(A(\xi_t^N) : t \geq 0)$ in $\mathcal{C}_1$ follows from the results of Section 4.5.

This in turn implies the C-tightness of $(\nu_t^N : t \geq 0)$ as cadlag Radon measure valued processes with the vague topology. Indeed, let $\varphi_k, k \in \mathbb{N}$ be a sequence of smooth functions from $\mathbb{R}$ to $[0,1]$ such that $\varphi_k(x)$ is 1 for $|x| \leq k$ and 0 for $|x| \geq k+1$. Then a diagonalization argument shows that C-tightness of $(\nu_t^N : t \geq 0)$ as cadlag Radon measure valued processes with the vague topology holds if and only if C-tightness of $(\varphi_k d\nu_t^N : t \geq 0)$ as cadlag $\mathcal{M}_F([-(k+1), k+1])$-valued processes with the weak topology holds. Here, $\mathcal{M}_F([-(k+1), k+1])$ denotes the space of finite measures on $[-(k+1), k+1]$. Now use Theorem II.4.1 in [11] to obtain C-tightness of $(\varphi_k d\nu_t^N : t \geq 0)$ in $D(\mathcal{M}_F([-(k+1), k+1]))$. The compact containment condition (i) in [11] is obvious. The second condition (ii) in [11] follows from (4.54) and the C-tightness of $(A(\xi_t^N) : t \geq 0)$ in $\mathcal{C}_1$ together with Lemma 4.4.4(a).

Observe in particular, that (4.54) implies the existence of a subsequence $\left(A(\xi_t^{N_k}), \nu_t^{N_k}\right)$ that converges to (u_t, ν_t). Hence, we can define variables with the same distributions on a different probability space such that with probability one, for all $T < \infty, \lambda > 0, \phi \in \mathcal{C}_c$,

$$\sup_{t \leq T} \left|\left| A(\xi_t^{N_k}) - u_t \right|\right|_{-\lambda} \to 0 \text{ as } k \to \infty,$$

$$\sup_{t \leq T} \left| < \phi, \nu_t^{N_k} > - < \phi, \nu_t > \right| \to 0 \text{ as } k \to \infty,$$

where we used $0 \leq A(\xi_t^{N_k}) \leq 1$ and thus $0 \leq u_t(x) \leq 1$ a.s. for the first limit. We obtain in particular

$$\nu_t(dx) = u_t(x)dx \text{ for all } t \geq 0.$$

It remains to investigate u_t in the special case, i.e. with no fixed kernel, i.e. where

$$q_{0j}^{(k,m,N)} = q_{1j}^{(k,m,N)}, j = 0, 1.$$

Take $\phi_t \equiv \phi \in \mathcal{C}_c^3$ in (4.33). We get

$$M_t^{(N)}(\phi) = < \nu_t^N, \phi > - < \nu_0^N, \phi > - \int_0^t < \nu_{s-}^N, \Delta(\phi) > ds - E_t^{(1)}(\phi) \quad (4.55)$$
$$- \sum_{k=0,1} (1-2k) \sum_{m\geq 2, j=0,1} q_{0j}^{(k,m,N)} \int_0^t < \left(F_j \circ A\left(\xi_{s-}^N\right)\right)$$
$$\times \left(F_{1-k} \circ A\left(\xi_{s-}^N\right)\right)^{m-1} \left(\delta_k \circ \xi_{s-}^N\right), \phi > ds - E_t^{(3)}(\phi).$$

From (4.27) and (4.32) and the Burkholder-Davis-Gundy inequality (4.45) we obtain that the error terms converge to zero for all $0 \leq t \leq T$ almost surely. Taylor's theorem further shows that (replace N_k by N for notational ease)

$$\Delta(\phi)(x_N) = \frac{N - \theta^{(N)}}{2c(N)N^{1/2}} \sum_{y \sim x_N} (\phi(y) - \phi(x_N))$$
$$= \frac{N - \theta^{(N)}}{c(N)N} \frac{\sqrt{N}}{2} \sum_{y \sim x_N} (\phi(y) - \phi(x_N))$$
$$\to \frac{\Delta\phi}{6}(x) \text{ as } x_N \to x \text{ and } N \to \infty$$

on the support of ϕ.

Using this in (4.55) we can show that $M_t^{(N)}(\phi)$ converges to a continuous martingale $M_t(\phi)$ satisfying

$$M_t(\phi) = \int \phi(x) u_t(x) dx - \int \phi(x) u_0(x) dx - \int_0^t \int \frac{\Delta\phi(x)}{6} u_s(x) dx ds \quad (4.56)$$
$$- \sum_{k=0,1} (1-2k) \sum_{m\geq 2, j=0,1} q_{0j}^{(k,m)} \int_0^t \int F_j(u_s(x))$$
$$\times (F_{1-k}(u_s(x)))^{m-1} F_k(u_s(x)) \phi(x) dx ds.$$

To exchange the limit in $N \to \infty$ with the infinite sum we used [12], Proposition 11.18 together with Hypothesis 4.2.20. Recall in particular, that $0 \leq F_l(u_s(x)) \leq 1$ for $l = 0, 1$. To show that $M_t(\phi)$ is indeed a martingale we used in particular (4.30) to see that $\langle M^{(N)}(\phi) \rangle_t \leq C(\lambda) t \parallel \phi \parallel_\lambda^2 < 1, e_{-2\lambda} >$ is uniformly bounded. Therefore, $(M_t^{(N)}(\phi) : N \geq N_0)$ and all its moments are uniformly integrable, using the Burkholder-Davis-Gundy inequality of the form (4.45) once more.

We can further calculate its quadratic variation by making use of (4.29) for $N \to \infty$ together with the uniform integrability of $((M_t^{(N)}(\phi))^2 : N \geq N_0)$.

Use our results for $\phi \in \mathcal{C}_c^3$, note that $\mathcal{C}_c^3$ is dense in $\mathcal{C}_c^2$ with respect to the norm $\parallel f \parallel \equiv \parallel f \parallel_\infty + \parallel f' \parallel_\infty + \parallel \Delta f \parallel_\infty$, and use (4.56) to see that u_t solves the martingale problem associated with the SPDE (4.18). It is now straightforward to show that, with respect to some white noise, u_t is actually a solution to (4.18) (see [13], V.20 for the similar argument in the case of SDEs).

4.7 Uniqueness in Law

To show uniqueness of all limit points of Section 4.6 in the case with no fixed kernel and with $< u_0, 1 >< \infty$, we need to show uniqueness in law of $[0,1]$-valued solutions to either (4.12) or (4.18) (recall Corollary 4.2.24). Indeed, as $0 \leq A(\xi_t^N)(x) \leq 1$ by definition, any limit point has to satisfy $u_t(x) \in [0,1]$.

We shall choose to prove weak uniqueness of (4.18), i.e. of

$$\begin{aligned}\frac{\partial u}{\partial t} =& \frac{\Delta u}{6} + \sum_{k=0,1} (1-2k) \sum_{m\geq 2, j=0,1} q_{0j}^{(k,m)} F_j(u) \left(F_{1-k}(u)\right)^{m-1} F_k(u) \\ &+ \sqrt{2u(1-u)}\dot{W} \\ =& \frac{\Delta u}{6} + u(1-u) \sum_{k=0,1} (1-2k) \sum_{m\geq 2, j=0,1} q_{0j}^{(k,m)} F_j(u) \left(F_{1-k}(u)\right)^{m-2} \\ &+ \sqrt{2u(1-u)}\dot{W} \\ \equiv& \frac{\Delta u}{6} + u(1-u)Q(u) + \sqrt{2u(1-u)}\dot{W}\end{aligned} \tag{4.57}$$

with initial condition u_0 in what follows. Observe that $|Q(u_s(x))| \leq C_Q$ with C_Q as in (4.43) because $0 \leq u_s(x) \leq 1$.

To check uniqueness in law of $[0,1]$-valued solutions we shall apply a version of Dawson's Girsanov Theorem, see Theorem IV.1.6 in [11], p. 252.

Let $\mathbb{P}^u$ denote the law of a solution to the SPDE (4.57) and $\mathbb{P}^v$ denote the unique law of the $[0,1]$-valued solution to the SPDE

$$\frac{\partial v}{\partial t} = \frac{\Delta v}{6} + \sqrt{2v(1-v)}\dot{W} \tag{4.58}$$

with $v_0 = u_0$. Reasons for existence and uniqueness of a $[0,1]$-valued solution to the latter can be found in Shiga [14], Example 5.2, p. 428. Note in particular that the solution v_t takes values in $\mathcal{C}_1$.

To prove weak uniqueness, we shall follow the reasoning of the proof of Theorem IV.1.6(a),(b) in [11] in a univariate setup. To follow the reasoning from [11], we need to show the following Lemma first.

Lemma 4.7.1. *Given $u_0 = v_0$ satisfying $< u_0, 1 >< \infty$, we have $\mathbb{P}^u$-a.s. $\int_0^t < u_s, 1 > ds < \infty$ for all $t \geq 0$ and $\mathbb{P}^v$-a.s. $\int_0^t < v_s, 1 > ds < \infty$ for all $t \geq 0$.*

Proof. We shall prove the claim for $\mathbb{P}^u$. The other claim then follows by considering the special case $Q \equiv 0$. As a first step we shall use a generalization of the weak form of (4.57) to functions in two variables. In the proof of Theorem 2.1

on p. 430 of [14] it is shown that for every $\psi \in \mathcal{D}^2_{rap}(T)$ and $0 < t < T$ we have

$$\begin{aligned} < u_t, \psi_t >= < u_0, \psi_0 > + \int_0^t < u_s, \left(\frac{\partial}{\partial s} + \frac{\Delta}{6}\right)\psi_s > ds \\ + \int_0^t < u_s(1-u_s)Q(u_s), \psi_s > ds \\ + \int_0^t \int \sqrt{2u_s(x)(1-u_s(x))}\psi_s(x) dW(x,s). \end{aligned} \tag{4.59}$$

Here we have for $T > 0$,

$$\begin{aligned} \mathcal{C}(\mathbb{R}) &= \{f : \mathbb{R} \to \mathbb{R} \text{ continuous}\}, \\ \mathcal{C}_{rap} &= \left\{ f \in \mathcal{C}(\mathbb{R}) \text{ s.t. } \sup_x e^{\lambda|x|}|f(x)| < \infty \text{ for all } \lambda > 0 \right\}, \\ \mathcal{C}^2_{rap} &= \left\{ \psi \in \mathcal{C}^2(\mathbb{R}) \text{ s.t. } \psi, \psi', \psi'' \in \mathcal{C}_{rap} \right\}, \\ \mathcal{D}^2_{rap}(T) &= \Big\{ \psi \in \mathcal{C}^{1,2}([0,T)\times\mathbb{R}) \text{ s.t. } \psi(t,\cdot) \text{ is } \mathcal{C}^2_{rap}\text{-valued continuous and} \\ &\qquad \frac{\partial \psi}{\partial t}(t,\cdot) \text{ is } \mathcal{C}_{rap}\text{-valued continuous in } 0 \le t < T \Big\}. \end{aligned}$$

Also observe that the condition (2.2) of [14] is satisfied as we have $0 \le u_s(x) \le 1$ and therefore $|Q(u_s(x))| \le C_Q$.

Now recall that the Brownian transition density is $p_s(x) = \frac{1}{\sqrt{2\pi s}} e^{-\frac{x^2}{2s}}$. Let $(P_s\phi)(x) = \int p_{\frac{s}{3}}(y-x)\phi(y)dy$ with $\phi \in \mathcal{C}_c^\infty, \phi \ge 0$ and let

$$\psi(s,x) = \psi_s(x) = e^{C_Q(T-s)}(P_{T-s}\phi)(x) \text{ and thus } \psi \in \mathcal{D}^2_{rap}(T).$$

Note that $\frac{\partial}{\partial s}(P_{T-s}\phi)(x) = -\frac{\Delta}{6}(P_{T-s}\phi)(x)$, where we used that $\frac{\partial}{\partial s}p_s(x) = \frac{1}{2}\Delta p_s(x)$. We obtain for the drift term in (4.59) that

$$\begin{aligned} &< u_s, \left(\frac{\partial}{\partial s} + \frac{\Delta}{6}\right)\psi_s > + < u_s(1-u_s)Q(u_s), \psi_s > \\ &=< u_s, -C_Q\psi_s - \frac{\Delta}{6}\psi_s + \frac{\Delta}{6}\psi_s > + < u_s(1-u_s)Q(u_s), \psi_s > \\ &\le 0 \end{aligned}$$

using that $\psi(s,x) \ge 0$ for $\phi \ge 0$. Additionally, the local martingale in (4.59) is a true martingale as

$$\begin{aligned} &\left\langle \int_0^\cdot \int \sqrt{2u_s(x)(1-u_s(x))}\psi_s(x)dW(x,s) \right\rangle_t \\ &= \int_0^t < 2u_s(1-u_s), \psi_s^2 > ds \le 2e^{2C_QT} \int_0^t < 1, (P_{T-s}\phi)^2 > ds \\ &\le 2e^{2C_QT} \|\phi\|_0 \int_0^t < 1, P_{T-s}\phi > ds = 2e^{2C_QT} \|\phi\|_0 < 1, \phi > t < \infty. \end{aligned}$$

Hence we obtain from (4.59) for all $0 < t < T$ after taking expectations

$$\mathbb{E}[< u_t, \psi_t >] \leq < u_0, \psi_0 >,$$

i.e.

$$e^{C_Q(T-t)}\mathbb{E}[< u_t, (P_{T-t}\phi) >] \leq e^{C_Q T} < u_0, (P_T\phi) > .$$

Now choose an increasing sequence of non-negative functions $\phi^n \in \mathcal{C}_c^\infty$ such that $\phi^n \uparrow 1$ for $n \to \infty$. Using the monotone convergence theorem, we obtain from the above

$$\begin{aligned} e^{C_Q(T-t)}\mathbb{E}[< u_t, 1 >] &= \lim_{n\to\infty} e^{C_Q(T-t)}\mathbb{E}[< u_t, (P_{T-t}\phi^n) >] \\ &\leq \lim_{n\to\infty} e^{C_Q T} < u_0, (P_T\phi^n) > = e^{C_Q T} < u_0, 1 > . \end{aligned}$$

Hence by the Fubini-Tonelli theorem,

$$\mathbb{E}\left[\int_0^t < u_s, 1 > ds\right] \leq < u_0, 1 > \int_0^t e^{C_Q s} ds < \infty$$

for all $t \geq 0$, which proves the claim. □

Lemma 4.7.2. *If $< u_0, 1 > < \infty$ the weak $[0,1]$-valued solution to (4.57) is unique in law. If we let*

$$\begin{aligned} R_t \equiv \exp\Bigg\{ &\int_0^t \int \frac{Q(v_s(x))}{2}\sqrt{2v_s(x)(1-v_s(x))}dW(x,s) \\ &-\frac{1}{2}\int_0^t \int \frac{(1-v_s(x))\,(Q(v_s(x)))^2}{2} v_s(x)dxds \Bigg\}, \end{aligned}$$

then

$$\left.\frac{d\mathbb{P}^u}{d\mathbb{P}^v}\right|_{\mathcal{F}_t} = R_t \text{ for all } t > 0, \tag{4.60}$$

where $\mathcal{F}_t$ is the canonical filtration of the process $v(t,x)$.

Proof. We proceed analogously to the proof of Theorem IV.1.6(a),(b) in [11]. Observe in particular that we take

$$T_n = \inf\left\{ t \geq 0 : \int_0^t \int \frac{(1-u_s(x))\,(Q(u_s(x)))^2}{2} u_s(x)dx + 1\; ds \geq n \right\}.$$

Lemma 4.7.1 shows that under $\mathbb{P}^u$

$$\int_0^t \int \frac{(1-u_s(x))\,(Q(u_s(x)))^2}{2} u_s(x)dxds \leq \frac{(C_Q)^2}{2}\int_0^t < u_s, 1 > ds < \infty$$

for all $t > 0$ $\mathbb{P}^u-$a.s. and so $T_n \uparrow \infty$ $\mathbb{P}^u$-a.s. As in Theorem IV.1.6(a) of [11] this gives uniqueness of the law $\mathbb{P}^u$ of a solution to (4.57). As in Theorem IV.1.6(b) of [11] the fact that $T_n \uparrow \infty$ $\mathbb{P}^v$-a.s. (from Lemma 4.7.1) shows that (4.60) defines a probability $\mathbb{P}^u$ which satisfies (4.57). □

Bibliography

[1] BURKHOLDER, D.L. Distribution function inequalities for martingales. *Ann. Probab.* (1973) **1**, 19–42. MR0365692

[2] COX, J.T. and DURRETT, R. and PERKINS, E.A. Rescaled voter models converge to super-Brownian motion. *Ann. Probab.* (2000) **28**, 185–234. MR1756003

[3] COX, J.T. and PERKINS, E.A. Rescaled Lotka-Volterra models converge to super-Brownian motion. *Ann. Probab.* (2005) **33**, 904–947. MR2135308

[4] COX, J.T. and PERKINS, E.A. Survival and coexistence in stochastic spatial Lotka-Volterra models. *Probab. Theory Related Fields* (2007) **139**, 89–142. MR2322693

[5] DAWSON, D.A. *Measure-valued Markov processes.* École d'éte de probabilités de Saint Flour, XXI (1991), 1–260, Lecture Notes in Math., 1541, Springer, Berlin, 1993. MR1242575

[6] DURRETT, R. *Ten lectures on particle systems.* Lectures on probability theory (Saint-Flour, 1993), 97–201, Lecture Notes in Math., 1608, Springer, Berlin, 1995. MR1383122

[7] LIGGETT, T.M. *Interacting Particle Systems*, Reprint of the 1985 original. Classics in Mathematics, Springer, Berlin, 2005. MR2108619

[8] LIGGETT, T.M. *Stochastic interacting systems: contact, voter and exclusion processes.* Springer, Berlin Heidelberg New York, 1999. MR1717346

[9] MUELLER, C. and TRIBE, R. Stochastic p.d.e.'s arising from the long range contact and long range voter processes. *Probab. Theory Related Fields* (1995) **102**, 519–545. MR1346264

[10] NEUHAUSER, C. and PACALA, S.W. An explicitly spatial version of the Lotka-Volterra model with interspecific competition. *Ann. Appl. Probab.* (1999) **9**, 1226–1259. MR1728561

[11] PERKINS, E.A. *Dawson-Watanabe superprocesses and measure-valued diffusions.* Lectures on Probability Theory and Statistics (Saint-Flour, 1999), 125–324, Lecture Notes in Math., 1781, Springer, Berlin, 2002. MR1915445

[12] ROYDEN, H.L. *Real Analysis*, Third edition. Macmillan Publishing Company, New York, 1988. MR1013117

[13] ROGERS, L.C.G. and WILLIAMS, D. *Diffusions, Markov Processes, and Martingales, vol. 2*, Reprint of the second (1994) edition. Cambridge Mathematical Univ. Press, Cambridge, 2000. MR1780932

[14] Shiga, T. Two contrasting properties of solutions for one-dimensional stochastic partial differential equations. *Canad. J. Math.* (1994) **46**, 415–437. MR1271224

[15] Walsh, J.B. *An Introduction to Stochastic Partial Differential Equations.* École d'éte de probabilités de Saint Flour, XIV (1984), 265–439, Lecture Notes in Math., 1180, Springer, Berlin, 1986. MR0876085

Chapter 5

Conclusion

The last three Chapters investigate different models for interacting multi-type populations from biology. All models under consideration are parameter-dependent and the behaviour of the parameters determines the behaviour of the system with respect to weak uniqueness or survival and coexistence results of types.

5.1 Overview of Results and Future Perspectives of the Manuscripts

5.1.1 Degenerate stochastic differential equations for catalytic branching networks

Chapter 2 establishes weak uniqueness for systems of SDEs, modelling catalytic branching networks. These networks can be obtained as limit points of branching particle systems. Weak uniqueness of the solutions to the SDE shows uniqueness of the limit points and hence implies convergence of the above approximations. It also makes certain additional tools available for analysis of the solution, as can be seen in the application of the results of Chapter 2 in Chapter 3. For instance, in the proof of existence of a stationary distribution for the normalized processes in Subsection A.2, weak uniqueness yields that the generator $\mathcal{A}$ satisfies the positive maximum principle.

This paper is an extension of Dawson and Perkins [6] (where the networks were essentially trees or cycles) to arbitrary networks. The additional dependency among catalysts led to a change of perspective from reactants to catalysts. In [6] every reactant j had one catalyst c_j only but as it turned out for networks it is more effective to consider every catalyst i with the set R_i of its reactants. In particular, the restriction from R_i to $\bar{R}_i$, including only reactants whose catalysts are all zero, turns out to be crucial. As a result, this paper introduces new ideas on how to handle networks where there exist catalytic interlinks between vertices.

As already mentioned in Subsection 1.1 of the introductory Chapter 1, the extension to networks becomes necessary for example in dimensions $d \geq 3$ in the renormalization analysis of hierarchically interacting multi-type branching models treated in Dawson, Greven, den Hollander, Sun and Swart [5]. The consideration of successive block averages leads to a renormalization transformation on the diffusion functions of a system of SDEs similar to (2.1), (2.2). Unfortunately, [5] only show preservation of the continuity of the coefficients of the SDE under this transformation but not preservation of Hölder-continuity. In [2], Bass and Perkins prove similar results to [6], i.e. they restrict themselves to networks with at most one catalyst per reactant, but drop the requirement of Hölder-continuity of the coefficients and replace it by continuity only. A future challenge would be to investigate how the ideas of my paper can be applied to extend the results of [2] to arbitrary networks. As a first step, [2] views the system of SDEs as a perturbation of a well-behaved system of SDEs with generator $\mathcal{A}^0$. As part of my paper I found an explicit representation of the semigroup P_t corresponding to $\mathcal{A}^0$ in the extension to the [6] setup. This representation directly carries over to the setting of [2]. The modification of the remaining reasoning in [2] to arbitrary networks remains to be done. Here we note that the extension to general graphs in Chapter 2 led to a number of technical problems in the approach of [6] even after the structure of the generator was resolved.

5.1.2 Long-term behaviour of a cyclic catalytic branching system

The results of Chapter 2 are used in Chapter 3 to establish weak uniqueness of the system of SDEs under consideration. The system involves catalytic branching and mutation between types.

Questions for survival and coexistence of types in the time-limit arise. Such questions naturally arise in biological competition models. Recall that Fleischmann and Xiong [7] investigated a cyclically catalytic super-Brownian motion. They showed global segregation (noncoexistence) of neighbouring types in the limit and other results on the finite time survival-extinction but they were not able to determine, if the overall sum dies out in the limit or not. I was able to show that in my setup the sum of all coordinates converges to zero but does not hit zero in finite time a.s. By changing my focus to the normalized processes I showed in particular that the normalized processes converge weakly to a unique stationary distribution that does not charge the set where at least one of the coordinates is zero.

The weakness of this manuscript lies in the restriction to constant positive coefficients γ^i and q_{ji} in the SDEs (3.1). It would be of interest to allow the coefficients to depend on the state of the system. As long as they are uniformly bounded away from zero and infinity, I conjecture a similar behaviour of the system but it would be of particular interest to see how the failure of uniform boundedness impacts questions on survival and coexistence. If one considers coefficients satisfying Hypothesis 2.1.2, the results of Chapter 2 can be applied to obtain weak uniqueness of the new system of SDEs as a first step.

Finally, following the approach in Section III.12.2 of [8], one could consider dilution flows instead of the normalized processes, that is, one removes the excess concentration so that the total concentration remains constant.

5.1.3 Convergence of rescaled competing species processes to a class of SPDEs

In Chapter 4, the tightness results obtained yield the relative compactness of the approximating particle systems. Limits along subsequences therefore exist for combinations of long-range kernel and fixed kernel interactions in the perturbations, where a wide class of admissible perturbations was found, including analytic functions as they appear in the spatial versions of Lotka-Volterra models of Neuhauser and Pacala [11]. It was particularly interesting to see that adding fixed kernel perturbations to the long-range case does not impact tightness.

In the long-range case I obtain that all subsequential limits satisfy a certain SPDE. It remains open to find the form of the limiting equations in the case of long-range dispersal in the presence of short-range, i.e. fixed kernel competition.

If I additionally assume finite initial mass in the long-range case, weak uniqueness of the limiting SPDE follows. As a consequence of this, the weak uniqueness of the limits of the approximating particle systems follows. It would be of interest to find necessary and sufficient conditions for weak uniqueness of the limiting SPDE. I conjecture $\int u_0(x)(1-u_0(x))dx < \infty$ to be a sufficient condition. If this condition is preserved by the dynamics it suffices for the Girsanov Lemma 4.7.2 to hold. In the special case under investigation in Mueller and Sowers [9] (see below), $\int u_t(x)(1-u_t(x))dx < \infty, \forall t > 0$ a.s. follows and so we expect this condition to define a state space for the solutions of the equations.

The class of SPDEs that I obtain as the limits of sequences of rescaled long-range perturbations of the voter model under some conditions on the parameters $\alpha_i^{(m)} \in \mathbb{R}, i = 0, 1, m \in \mathbb{N}$ and their counterparts in the approximating models are

$$\frac{\partial u}{\partial t} = \frac{\Delta u}{6} + (1-u)u \left\{ \sum_{m=0}^{\infty} \alpha_0^{(m+1)} u^m - \sum_{m=0}^{\infty} \alpha_1^{(m+1)} (1-u)^m \right\} + \sqrt{2u(1-u)}\dot{W}.$$

As I work at the critical range $\sqrt{N}$, while Cox and Perkins [3] works with longer range interactions, I obtain a wide class of non-linear drifts. This opens up the possibility to interpret the limiting SPDEs and their behaviour via their approximating long-range particle systems and vice versa. For instance, a future challenge would be to use properties of the SPDE to obtain results on the approximating particle systems, following the ideas of [3] and Cox and Perkins [4]. As an example, recall Remark 4.2.15 where I obtained the SPDE

$$\frac{\partial u}{\partial t} = \frac{\Delta u}{6} + (1-u)u \left\{ \lambda' - \alpha_{10} + u\left(\alpha_{01} + \alpha_{10}\right) \right\} + \sqrt{2u(1-u)}\dot{W}$$

with parameters $\lambda', \alpha_{10}, \alpha_{01} \in \mathbb{R}$ as the limit of spatial versions of the Lotka-Volterra model with competition and fecundity parameters near one. We can rewrite this SPDE to

$$\frac{\partial u}{\partial t} = \frac{\Delta u}{6} + \theta_0(1-u)u^2 - \theta_1 u(1-u)^2 + \sqrt{2u(1-u)}\dot{W}, \tag{5.1}$$
$$u(0,x) \equiv u_0(x) \geq 0,$$

where $\theta_i \in \mathbb{R}, i = 0, 1$. For $\theta_1 = -\theta_0 < 0$ one obtains after rescaling the Kolmogorov-Petrovskii-Piscuinov (KPP) equation driven by Fisher-Wright noise. This SPDE has already been investigated in Mueller and Sowers [9] in detail, where the existence of travelling waves was shown for θ_0 big enough.

A major question is how the change in the drift, in particular, the possibly additional zero at $\frac{\theta_1}{\theta_0+\theta_1} \in (0,1)$, impacts the set of parameters for survival (i.e. $\limsup_{t\to\infty} < u_t, 1 >> 0$ with positive probability), coexistence (i.e. there exists a stationary distribution giving zero mass to the configurations 0 and 1) and extinction (i.e. $\limsup_{t\to\infty} < u_t, 1 >= 0$ with probability 1) and if there exist phase transitions. Aronson and Weinberger [1] showed for instance in Corollary 3.1(ii) that for $\theta_0 < 0, \theta_1 < 0$, the corresponding deterministic PDE converges to the intermediate zero $\frac{\theta_1}{\theta_0+\theta_1}$ of the drift term uniformly on bounded sets if $u_0 \not\equiv 0, 1$.

The author conjectures that there are parameter regions for (5.1) that yield survival and others that yield extinction. To prove survival the author envisions to apply methods of Mueller and Tribe [10] to the SPDE (5.1). In [10] and Tribe [12], rescaled versions of the SPDE

$$\frac{\partial u}{\partial t} = \frac{\Delta u}{6} + \theta u - u^2 + \sqrt{u}\dot{W},\ \theta > 0 \tag{5.2}$$

were under investigation. Results on the existence of a phase transition between extinction and survival in terms of θ are obtained in the former paper and the existence of travelling wave solutions in the latter. Unfortunately, the proof of extinction in [10] and the proof of existence of travelling waves in [12] relies on the additive properties of the fluctuation term in (5.2) (also recall the discussion of additive properties of super-Brownian motion in the beginning of the introductory Chapter 1), which makes the application of their methods to SPDEs of the form (5.1) difficult. On the other hand, [10] shows that for θ large, the drift term of the SPDE (5.2) outcompetes the fluctuation term. The proof for survival then uses a comparison of $u(t,x)$ to oriented site percolation to prove survival for θ big. It should be possible to apply their reasoning together with the results of [1] for the corresponding deterministic PDE to the SPDE (5.1) to show survival of types in certain parameter-regions. Extinction on the other hand seems much more delicate to prove.

Bibliography

[1] Aronson, D.G. and Weinberger, H.F. Multidimensional Nonlinear Diffusion Arising in Population Genetics. *Adv. Math.* (1978) **30**, 33–76. MR0511740

[2] Bass, R.F. and Perkins, E.A. Degenerate stochastic differential equations arising from catalytic branching networks. *Electron. J. Probab.* (2008) **13**, 1808–1885. MR2448130

[3] Cox, J.T. and Perkins, E.A. Rescaled Lotka-Volterra Models converge to Super-Brownian Motion. *Ann. Probab.* (2005) **33**, 904–947. MR2135308

[4] Cox, J.T. and Perkins, E.A. Survival and coexistence in stochastic spatial Lotka-Volterra models. *Probab. Theory Related Fields* (2007) **139**, 89–142. MR2322693

[5] Dawson, D.A. and Greven, A. and den Hollander, F. and Sun, R. and Swart, J.M. The renormalization transformation for two-type branching models. *Ann. Inst. H. Poincaré Probab. Statist.* (2008) **44**, 1038–1077. MR2469334

[6] Dawson, D.A. and Perkins, E.A. On the uniqueness problem for catalytic branching networks and other singular diffusions. *Illinois J. Math.* (2006) **50**, 323–383 (electronic). MR2247832

[7] Fleischmann, K. and Xiong, J. A cyclically catalytic super-brownian motion. *Ann. Probab.* (2001) **29**, 820–861. MR1849179

[8] Hofbauer, J. and Sigmund, K. *The Theory of Evolution and Dynamical Systems*. London Math. Soc. Stud. Texts, vol. 7, Cambridge Univ. Press, Cambridge, 1988. MR1071180

[9] Mueller, C. and Sowers, R.B. Random Travelling Waves for the KPP Equation with Noise. *J. Funct. Anal.* (1995) **128**, 439–498. MR1319963

[10] Mueller, C. and Tribe, R. A phase transition for a stochastic PDE related to the contact process. *Probab. Theory Related Fields* (1994) **100**, 131–156. MR1296425

[11] Neuhauser, C. and Pacala, S.W. An explicitly spatial version of the Lotka-Volterra model with interspecific competition. *Ann. Appl. Probab.* (1999) **9**, 1226–1259. MR1728561

[12] Tribe, R. A travelling wave solution to the Kolmogorov equation with noise. *Stochastics Stochastics Rep.* (1996) **56**, 317–340. MR1396765

Appendix A

Appendix for Chapter 3

A.1 $\tilde{a}$ is non-singular

Corollary A.1.1. *The matrix $\tilde{a}$ is non-singular for all $x \in \tilde{S}$, where*

$$\tilde{S} = \left([0,1]^{d-1} \cap \left\{\sum_{i=1}^{d-1} x_i \le 1\right\}\right) \setminus \left\{x : \exists i : x_i = 0 \ or \ \sum_{i=1}^{d-1} x_i = 1\right\}.$$

Proof. Recall that $\sigma \in \mathcal{M}(d,d)$ (the space of $d \times d$-matrices) and $a = \sigma\sigma^T \in \mathcal{M}(d,d)$. Let $\bar{\sigma} \in \mathcal{M}(d-1,d)$ be constructed from σ by deleting the last line of the matrix (i.e. by deleting the last equation for Y_t^d of our system of SDEs). Then $\tilde{a} = \bar{\sigma}\bar{\sigma}^T \in \mathcal{M}(d-1,d-1)$. Further let $\tilde{\sigma} \in \mathcal{M}(d-1,d-1)$ be the matrix obtained from $\bar{\sigma}$ by deleting the last column.

We claim that if $\tilde{\sigma}$ is non-singular, then $\tilde{a}$ is non-singular as well. Indeed, let $v \in \mathcal{M}(d-1,1)$ denote the last column of $\bar{\sigma}$ and suppose $\tilde{\sigma}$ is non-singular, then

$$\begin{aligned}\det(\tilde{a}) &= \det(\bar{\sigma}\bar{\sigma}^T) = \det(\tilde{\sigma}\tilde{\sigma}^T + vv^T) = \det\big(\tilde{\sigma}\tilde{\sigma}^T\big)\left(1 + v^T(\tilde{\sigma}\tilde{\sigma}^T)^{-1}v\right) \\ &= \det\big(\tilde{\sigma}\tilde{\sigma}^T\big)\left(1 + \parallel \tilde{\sigma}^{-1}v \parallel^2\right) = (\det(\tilde{\sigma}))^2\left(1 + \parallel \tilde{\sigma}^{-1}v \parallel^2\right) > 0.\end{aligned}$$

Recall that for $i,j \in \{1,\dots,d-1\}$ we have $\tilde{\sigma}_{ii}(x) = (1-x_i)\sqrt{2\gamma^i x_i x_{i+1}}$ and $\tilde{\sigma}_{ij}(x) = -x^i\sqrt{2\gamma^j x_j x_{j+1}}$ if $i \neq j$, where we set $x_d \equiv 1 - \sum_{i=1}^{d-1} x_i$. Suppose that $x_i > 0$ for all $i \in \{1,\dots,d-1\}$ and $\sum_{i=1}^{d-1} x_i < 1$. We shall show that in this case $\tilde{\sigma}$ is non-singular. As a first step we divide the i^{th} line of $\tilde{\sigma}$ by x_i for $i = 1,\dots,d-1$. We obtain

$$\begin{aligned}\frac{\det(\tilde{\sigma}(x))}{\prod_{i=1}^{d-1} x_i} &= \det\begin{pmatrix} d_1 & a_2 & a_3 & \dots & a_{d-1} \\ a_1 & d_2 & a_3 & \dots & a_{d-1} \\ a_1 & a_2 & d_3 & \dots & a_{d-1} \\ \dots & \dots & \dots & \dots & \dots \\ a_1 & a_2 & a_3 & \dots & d_{d-1}\end{pmatrix} \\ &= \det\begin{pmatrix} d_1 - a_1 & a_2 - d_2 & 0 & \dots & 0 & 0 \\ 0 & d_2 - a_2 & a_3 - d_3 & \dots & 0 & 0 \\ 0 & 0 & d_3 - a_3 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & d_{d-2} - a_{d-2} & a_{d-1} - d_{d-1} \\ a_1 & a_2 & a_3 & \dots & a_{d-2} & d_{d-1}\end{pmatrix} \\ &\equiv \det A_1,\end{aligned}$$

where we used the line operations $i \to i-(i+1)$ on all but the last line and set

$$d_i - a_i = (1-x_i)\sqrt{2\gamma^i \frac{x_{i+1}}{x_i}} + \sqrt{2\gamma^i x_i x_{i+1}} = \sqrt{2\gamma^i \frac{x_{i+1}}{x_i}} = -\frac{a_i}{x_i}. \tag{A.1}$$

Based on the first column we can calculate the determinant of $\tilde{\sigma}$.

$$\det(A_1) \equiv (d_1 - a_1)\det(A_2) + (-1)^d a_1 \prod_{i=2}^{d-1}(a_i - d_i), \tag{A.2}$$

where we obtain the matrix A_2 by crossing out the first row and column of A_1. We obtain recursively for $k = 1, \dots, d-3$, using (A.1) that

$$\det(A_k) = -\frac{a_k}{x_k}\det(A_{k+1}) + a_k(-1)^{d-k+1}\prod_{i=k+1}^{d-1}\frac{a_i}{x_i}. \tag{A.3}$$

By using (A.3) recursively in (A.2) we get

$$\det(A_1) = (-1)^{d-3}\prod_{i=1}^{d-3}\frac{a_i}{x_i}\det(A_{d-2}) + (-1)^d \prod_{i=1}^{d-1} a_i \sum_{i=1}^{d-3}\prod_{j=1,\dots,d-1,j\neq i}\frac{1}{x_j}.$$

Finally,

$$\det(A_{d-2}) = \det\begin{pmatrix} d_{d-2} - a_{d-2} & a_{d-1} - d_{d-1} \\ a_{d-2} & d_{d-1}\end{pmatrix} \overset{(A.1)}{=} -\frac{a_{d-2}d_{d-1}}{x_{d-2}} - \frac{a_{d-2}a_{d-1}}{x_{d-1}}$$

and thus (recall that $x_i \neq 0$ and thus $a_i \neq 0$ for $i = 1, \dots, d$)

$$\begin{aligned}
&\det(\tilde{\sigma}(x)) = 0 \iff \det(A_1) = 0 \\
&\iff \sum_{i=1}^{d-2}\prod_{j=1,\dots,d-1,j\neq i}\frac{1}{x_j} + \frac{d_{d-1}}{a_{d-1}}\prod_{j=1,\dots,d-1,j\neq d-1}\frac{1}{x_j} = 0 \\
&\iff \sum_{i=1}^{d-2}\prod_{j=1,\dots,d-1,j\neq i}\frac{1}{x_j} - \frac{1-x_{d-1}}{x_{d-1}}\prod_{j=1,\dots,d-1,j\neq d-1}\frac{1}{x_j} = 0 \\
&\iff \sum_{i=1}^{d-2} x_i - (1-x_{d-1}) = 0 \iff \sum_{i=1}^{d-1} x_i = x_d = 1,
\end{aligned}$$

which is a contradiction to $x \in \tilde{S}$. Hence $\det(\tilde{\sigma}(x)) \neq 0$ for all $x \in \tilde{S}$. □

A.2 Proof of Proposition 3.2.18

Proof. The main part of the proof is taken from Dawson, Greven, den Hollander, Sun and Swart [1], Section 3.1 and adjusted to our setting.

Existence. Denote the distribution of $Y_t \in [0,1]^d$ by μ_t, with $\mu_0 = \delta_y$ for some arbitrary $y \in S$. As the state space $[0,1]^d$ is compact, $\{\nu_t : \nu_t \equiv \frac{1}{t}\int_0^t \mu_s ds\}_{t\geq 0}$ forms a tight family of distributions.

In this case Theorem III.2.2 in Ethier and Kurtz [2] implies that $\{\nu_t\}_{t\geq 0}$ is relatively compact for the Prohorov metric. As $[0,1]^d$ is Polish, Theorem III.1.7 in [2] gives the existence of a limit. Also note that the convergence of a sequence of $\{\nu_t\}_{t\geq 0}$ is equivalent to weak convergence by Theorem III.3.1 of [2].

Taking this together we find a sequence (t_n) tending to infinity such that ν_{t_n} converges weakly to a limiting distribution ν. The goal will be now to apply Theorem IV.9.17 of [2], where $\hat{\mathcal{C}}([0,1]^d) = \mathcal{C}([0,1]^d)$ by definition (cf. definition before Lemma IV.2.1) due to the compactness of our state space.

To this purpose note first that the generator corresponding to (3.15) is given by

$$\mathcal{A}f(x) = \sum_{i=1}^d b_i(x)\partial_i f(x) + \frac{1}{2}\sum_{i,j=1}^d a_{ij}(x)\partial_{ij} f(x),$$

where $a = \sigma\sigma^T$ and b are as in (3.19) and $\mathcal{D}(\mathcal{A}) = \mathcal{C}^2([0,1]^d)$.

Note further that our system of SDEs has a solution that is unique in law (cf. Proposition 3.2.13). Hence Theorem V.(21.2) and Remark V.(21.9) in Rogers and Williams [3] yield that the solution is strong Markov. Hence $\mathcal{A}$ is the generator of a strong Markov process and thus satisties the positive maximum principle by [2], Theorem IV.2.2.

Moreover, if ν is the limit of ν_{t_n} then for any $f \in \mathcal{C}^2([0,1]^d)$ we have $\mathcal{A}f \in \mathcal{C}([0,1]^d)$ and

$$\begin{aligned}
\mathbb{E}_\nu[(\mathcal{A}f)(Y_0)] &= \lim_{n\to\infty} \mathbb{E}_{\nu_{t_n}}[(\mathcal{A}f)(Y_0)] \\
&= \lim_{n\to\infty} \frac{1}{t_n}\int_0^{t_n} \mathbb{E}_{\mu_s}[(\mathcal{A}f)(Y_0)]\, ds \\
&= \lim_{n\to\infty} \frac{1}{t_n}\int_0^{t_n} \mathbb{E}_{\mu_0}[(\mathcal{A}f)(Y_s)]\, ds \\
&= \lim_{n\to\infty} \frac{1}{t_n}\mathbb{E}_{\mu_0}\left[\int_0^{t_n} (\mathcal{A}f)(Y_s) ds\right] \\
&= \lim_{n\to\infty} \frac{1}{t_n}\mathbb{E}_{\mu_0}[f(Y_{t_n}) - f(Y_0)] \\
&= 0.
\end{aligned}$$

Here we used $\nu_{t_n} \Rightarrow \nu$ for $\mathcal{A}f \in \mathcal{C}([0,1]^d)$ in the first equality, the definition of ν_t in the second, $\mathcal{A}f \in \mathcal{C}([0,1]^d)$ in the fourth equality, that Y solves the martingale problem for $\mathcal{A}$ in the fifth equality and f bounded in the last.

Finally observe that $\mathcal{C}^2([0,1]^d)$ forms an algebra of functions dense in $\mathcal{C}([0,1]^d)$. Taken the above together, we can apply Theorem IV.9.17 of [2] and obtain the existence of a stationary solution.

Uniqueness. Let Y_t be the unique strong Markov solution to (3.15). Recall that we already showed that every equilibrium distribution for $\mathcal{A}$ doesn't put mass on $\mathcal{N} = \{y : \exists i : y_i = 0\}$ in Proposition 3.2.17 and that $\sum_i Y_t^i = 1$ for all $t \geq 0$ in Proposition 3.2.13. Hence, if Y_t has two distinct equilibrium distributions, concentrated on $[0,1]^d$, or to be more precise, on

$$S \equiv \left([0,1]^d \backslash \partial [0,1]^d\right) \cap \left\{ y : \sum_i y_i = 1 \right\},$$

then we can find two extremal equilibrium distributions μ and ν that are singular with respect to each other (see for instance Exercise 6.9 in Varadhan [5]). As $\mu(B) = \mathbb{P}_\mu(Y_t \in B) = \int p_t(x,B)d\mu(x)$, there have to exist $x, y \in S$ such that the transition kernels $p_t(x,dz)$ and $p_t(y,dz)$ are mutually singular for all $t > 0$ as well. Also, as μ respectively ν do not put mass on $\mathcal{N}$ the same holds for $p_t(x,dz)$ respectively $p_t(y,dz)$.

In what follows we shall consider the process $\tilde{Y}_t \equiv (Y_t^1, \ldots, Y_t^{d-1}) \in [0,1]^{d-1}$ with transition kernels $\tilde{p}_t$ instead. The martingale problem for the resulting SDE for $\tilde{Y}$ is consequently well-posed as the corresponding martingale problem for Y is well-posed.

Let $\tilde{p}_t(\tilde{x}, d\tilde{z})$ and $\tilde{p}_t(\tilde{y}, d\tilde{z})$ be the resulting transition kernels corresponding to $p_t(x,dz)$ and $p_t(y,dz)$.

For

$$\begin{aligned} \tilde{x} \in \tilde{S} &\equiv \left([0,1]^{d-1} \cap \left\{ \sum_{i=1}^{d-1} \tilde{x}_i \leq 1 \right\} \right) \setminus \left\{ \tilde{x} : \exists i : \tilde{x}_i = 0 \text{ or } \sum_{i=1}^{d-1} \tilde{x}_i = 1 \right\} \\ &= \left([0,1]^{d-1} \backslash \partial [0,1]^{d-1}\right) \cap \left\{ \tilde{x} : 0 < \sum_{i=1}^{d-1} \tilde{x}_i < 1 \right\}, \end{aligned}$$

$\tilde{a}_{ij}(\tilde{x})$ is non-singular by Corollary A.1.1 of the Appendix. Now we can appeal to Theorem B.4 ("Support theorem for uniformly elliptic diffusions") of [1] with $\tilde{x}, \tilde{y} \in D$, $\bar{D} \subset \tilde{S}$, where D is an arbitrarily fixed ball. Here observe that $\tilde{S}$ is an open subset of $[0,1]^{d-1}$. The Theorem allows us to transport the diffusions started at $\tilde{x}$ respectively $\tilde{y}$ to a common small neighbourhood with positive probability. Subsequently we can apply Corollary B.3 ("Transition density for diffusions restricted to bounded domains") of [1] to see that $\tilde{p}_t(\tilde{x}, d\tilde{z})$ and $\tilde{p}_t(\tilde{y}, d\tilde{z})$ and hence $p_t(x,dz)$ and $p_t(y,dz)$ cannot be mutually singular for all $t > 0$.

Convergence. Once more, we shall consider the process $\tilde{Y}$ instead of Y and the state space $\tilde{S}$ instead of S. Let $\tilde{\pi}$ be the unique equilibrium distribution of $\tilde{Y}$ corresponding to the unique equilibrium distribution π of Y.

Firstly, note that by Theorem B.4 of [1] the equilibrium distribution $\tilde{\pi}$ assigns positive measure to every open subset of $\tilde{S}$.

Secondly we shall show that $\mathcal{L}(\tilde{Y}_t|\tilde{Y}_0 = x) \Rightarrow \tilde{\pi}$ for all $x \in \tilde{S}$ in three steps, namely by showing first that it holds for almost all $x \in \tilde{S}$ w.r.t. $\tilde{\pi}$. Then we shall extend this result to Lebesgue almost every $x \in \tilde{S}$ and finally we shall conclude that this implies that it holds for all $x \in \tilde{S}$.

To prove the first step, we first choose $x^* \in \tilde{S}$ arbitrary but fixed and let D be open such that $x^* \in D$, $\bar{D} \subset \tilde{S}$. Recall that $\tilde{a}$ is non-singular on $\tilde{S}$ and $\tilde{S}$ is an open subset of $[0,1]^{d-1}$. Let $\tilde{Y}_t, \tilde{Z}_t$ be two independent copies of the process on $[0,1]^{d-1}$. Then the joint process $(\tilde{Y}_t, \tilde{Z}_t)$ is strong Markov and has a unique equilibrium distribution given by the product measure $\tilde{\pi} \otimes \tilde{\pi}$. By Theorem 6.9 in [5] and the following Remarks the process $(\tilde{Y}_t, \tilde{Z}_t)$ started in equilibrium, i.e. with $\mathcal{L}((\tilde{Y}_t, \tilde{Z}_t)) = \tilde{\pi} \otimes \tilde{\pi}$ is ergodic.

As the equilibrium distribution $\tilde{\pi}$ assigns positive measure to every open subset of $\tilde{S}$, $\tilde{\pi} \otimes \tilde{\pi}$ assigns positive measure to $B_\epsilon(x^*) \times B_\epsilon(x^*) \subset D \times D$ for ϵ small enough. Therefore $(\tilde{Y}_t, \tilde{Z}_t)$ visits the set $B_\epsilon(x^*) \times B_\epsilon(x^*)$ after any finite time T a.s. by the ergodic theorem. We obtain in particular that for almost all (x, x') w.r.t. $\tilde{\pi} \otimes \tilde{\pi}$, $(\tilde{Y}_t, \tilde{Z}_t)$ started at (x, x') visits $B_\epsilon(x^*) \times B_\epsilon(x^*)$ after any finite time T a.s. Fix such a (x, x').

In what follows we shall start two independent processes $\tilde{Y}_t$ and $\tilde{Z}_t$ with initial conditions x respectively x' as above and denote their laws by $\mathbb{P}^x$ respectively $\mathbb{P}^{x'}$.

Let the first exit time from D be

$$\tau_D(\omega) \equiv \inf\{t \geq 0 : \omega(t) \notin D\}.$$

By Corollary B.3, for each $\delta > 0$ and $z \in D$, the measure

$$\mu_\delta^D(z, \cdot) \equiv \mathbb{P}^z(\omega : \delta < \tau_D(\omega), \omega(\delta) \in \cdot)$$

admits a density $p_\delta^D(z, \cdot)$ with respect to Lebesgue measure. Moreover, (B.2) of the Corollary yields that for ϵ, δ sufficiently small we have uniformly for $y, y' \in \overline{B_\epsilon(x^*)}$,

$$\int p_\delta^D(y, u) \wedge p_\delta^D(y', u) du \geq \frac{1}{2}, \tag{A.4}$$

where we used that $a + b - (a \vee b - a \wedge b) = a \vee b + a \wedge b - (a \vee b - a \wedge b) \geq 1 \iff a \wedge b \geq 1/2$. We obtain in particular that for $y \in \overline{B_\epsilon(x^*)}$ (and analogously for $y' \in \overline{B_\epsilon(x^*)}$)

$$\begin{aligned} \mathbb{P}^y(\omega : \omega(\delta) \in du) &\geq \mathbb{P}^y(\omega : \delta < \tau_D(\omega), \omega(\delta) \in du) \\ &= p_\delta^D(y, u) du \geq p_\delta^D(y, u) \wedge p_\delta^D(y', u) du. \end{aligned} \tag{A.5}$$

For $y, y' \in \overline{B_\epsilon(x^*)}$ fixed let $\mu^1_{(y,y')}$ be the measure on $[0,1]^{d-1} \times [0,1]^{d-1}$ defined by

$$\mu^1_{(y,y')}(A \times B) \equiv \int_{A \times B} \left(p_\delta^D(y, u) \wedge p_\delta^D(y', u)\right) \left(p_\delta^D(y, v) \wedge p_\delta^D(y', v)\right) du dv$$

for $A, B \in \mathcal{B}([0,1]^{d-1})$ and observe that

$$\mu^1_{(y,y')}(A \times B) = \mu^1_{(y,y')}(B \times A). \tag{A.6}$$

Let

$$\mu^2_{(y,y')}(A \times B) \equiv \mathbb{P}^y(\omega : \omega(\delta) \in A)\mathbb{P}^{y'}(\omega : \omega(\delta) \in B) - \mu^1_{(y,y')}(A \times B),$$

which is non-negative by (A.5) and note that by (A.4),

$$\mu^2_{(y,y')}\big([0,1]^{d-1} \times [0,1]^{d-1}\big) \leq \frac{3}{4} \text{ for all } y, y' \in \overline{B_\epsilon(x^*)}. \tag{A.7}$$

In what follows we shall give the motivation for the later rigorous definitions and calculations.

We obtain that whenever $(\tilde{Y}_t, \tilde{Z}_t)$, starting at (x, x') enters $\overline{B_\epsilon(x^*) \times B_\epsilon(x^*)}$ at a random time, say T_1 and through some random point $(y, y') \in \overline{B_\epsilon(x^*) \times B_\epsilon(x^*)}$ we can decompose the conditional law of $(\tilde{Y}_t, \tilde{Z}_t)$ at time $T_1 + \delta$ as follows:

$$\mathbb{P}^y(\omega : \omega(\delta) \in A)\mathbb{P}^{y'}(\omega : \omega(\delta) \in B) = \mu^1_{(y,y')}(A \times B) + \mu^2_{(y,y')}(A \times B). \tag{A.8}$$

Now we shall successively decompose the law of the process $(\tilde{Y}_t, \tilde{Z}_t)$ at times $T_k + \delta, k \in \mathbb{N}$, where T_k is the first time after $T_{k-1} + \delta$ that $(\tilde{Y}_t, \tilde{Z}_t)$ enters $\overline{B_\epsilon(x^*) \times B_\epsilon(x^*)}$ (see (A.9) below). For instance, if at time T_1, $(\tilde{Y}_t, \tilde{Z}_t)$ enters $B_\epsilon(x^*) \times B_\epsilon(x^*)$ through some random point $(y, y') \in \overline{B_\epsilon(x^*) \times B_\epsilon(x^*)}$, we decompose the law of $(\tilde{Y}_t, \tilde{Z}_t)$ at time $T_1 + \delta$ into $\mu^1_{(y,y')}, \mu^2_{(y,y')}$ as above. Next we consider $(\tilde{Y}_{T_1+\delta+\cdot}, \tilde{Z}_{T_1+\delta+\cdot})$, starting in $\mu^1_{(y,y')}$ resp. $\mu^2_{(y,y')}$. As $\mu^1_{(y,y')}$ yields a common part, we do not decompose the law of $(\tilde{Y}_{T_1+\delta+\cdot}, \tilde{Z}_{T_1+\delta+\cdot})$ starting in $\mu^1_{(y,y')}$ any further. On the other hand, starting in $\mu^2_{(y,y')}$, we wait until $(\tilde{Y}_{T_1+\delta+\cdot}, \tilde{Z}_{T_1+\delta+\cdot})$ enters $\overline{B_\epsilon(x^*) \times B_\epsilon(x^*)}$ again, say at time $T_1' = T_2 - (T_1 + \delta)$ through the point $(u, u') \in \overline{B_\epsilon(x^*) \times B_\epsilon(x^*)}$. T_1' is finite $\mu^2_{(y,y')}$-a.s. as $\mu^2_{(y,y')}$ is absolutely continuous with respect to $\mathbb{P}^y(\omega : \omega(\delta) \in \cdot) \otimes \mathbb{P}^{y'}(\omega : \omega(\delta) \in \cdot)$ by (A.8).

Here recall that (y, y') was the first entrance point of $B_\epsilon(x^*) \times B_\epsilon(x^*)$ by $(\tilde{Y}_t, \tilde{Z}_t)$ under $\mathbb{P}^x \otimes \mathbb{P}^{x'}$ and that $(\tilde{Y}_t, \tilde{Z}_t)$ started at (x, x') visits $B_\epsilon(x^*) \times B_\epsilon(x^*)$ after any finite time T a.s. Hence $(\tilde{Y}_t, \tilde{Z}_t)$ started at (x, x') visits $B_\epsilon(x^*) \times B_\epsilon(x^*)$ again at some finite random time $T_2 = T_1 + \delta + T_1'$. Thus, starting in $\mu^2_{(y,y')}$, we can decompose $(\tilde{Y}_{T_1+\delta+T_1'+\delta}, \tilde{Z}_{T_1+\delta+T_1'+\delta})$ in $\mu^1_{(u,u')}$ and $\mu^2_{(u,u')}$. Now iterate the above.

To be more explicit, let $U_0 = 0$ and define stopping times

$$T_k = \inf\Big\{t \geq U_{k-1} : \Big(\tilde{Y}_t, \tilde{Z}_t\Big) \in B_\epsilon(x^*) \times B_\epsilon(x^*)\Big\} \text{ and } U_k = T_k + \delta \tag{A.9}$$

for $k \in \mathbb{N}$ and $\delta > 0$. Then almost surely $T_k < \infty$ for all $k \in \mathbb{N}$.

By the strong Markov property of the process $(\tilde{Y}_t, \tilde{Z}_t)$ we can condition on $\mathcal{F}_{U_1}$ and obtain for $n \in \mathbb{N}$ arbitrarily fixed,

$$\begin{aligned}
&\Big(\mathbb{P}^x \otimes \mathbb{P}^{x'}\Big)\Big(\tilde{Y}_t \in A, \tilde{Z}_t \in B\Big) \\
&= \Big(\mathbb{P}^x \otimes \mathbb{P}^{x'}\Big)\Big(\tilde{Y}_t \in A, \tilde{Z}_t \in B, t < U_n\Big) \\
&\quad + \Big(\mathbb{P}^x \otimes \mathbb{P}^{x'}\Big)\Big[\Big(\mathbb{P}^{\tilde{Y}_{T_1(\omega)}} \otimes \mathbb{P}^{\tilde{Z}_{T_1(\omega)}}\Big)\Big(\tilde{Y}_{t-T_1(\omega)} \in A, \tilde{Z}_{t-T_1(\omega)} \in B, \\
&\quad t \geq U_1(\omega) + U_{n-1} \circ \theta_\delta\Big) 1(t \geq U_1(\omega))\Big].
\end{aligned}$$

Here θ_δ denotes the shift-operator $\theta_\delta(\omega(\cdot)) = \omega(\delta + \cdot)$.

Using (A.8) we can rewrite this as

$$\begin{aligned}
&\Big(\mathbb{P}^x \otimes \mathbb{P}^{x'}\Big)\Big(\tilde{Y}_t \in A, \tilde{Z}_t \in B\Big) \qquad \text{(A.10)} \\
&= \Big(\mathbb{P}^x \otimes \mathbb{P}^{x'}\Big)\Big(\tilde{Y}_t \in A, \tilde{Z}_t \in B, t < U_n\Big) \\
&\quad + \Big(\mathbb{P}^x \otimes \mathbb{P}^{x'}\Big)\Big[\mu^1_{(\tilde{Y}_{T_1(\omega)}, \tilde{Z}_{T_1(\omega)})}(dw, dw')\Big(\mathbb{P}^w \otimes \mathbb{P}^{w'}\Big)\Big(\tilde{Y}_{t-U_1(\omega)} \in A, \\
&\quad \tilde{Z}_{t-U_1(\omega)} \in B, t \geq U_1(\omega) + U_{n-1}\Big) 1(t \geq U_1(\omega))\Big] \\
&\quad + \Big(\mathbb{P}^x \otimes \mathbb{P}^{x'}\Big)\Big[\mu^2_{(\tilde{Y}_{T_1(\omega)}, \tilde{Z}_{T_1(\omega)})}(dw, dw')\Big(\mathbb{P}^w \otimes \mathbb{P}^{w'}\Big)\Big(\tilde{Y}_{t-U_1(\omega)} \in A, \\
&\quad \tilde{Z}_{t-U_1(\omega)} \in B, t \geq U_1(\omega) + U_{n-1}\Big) 1(t \geq U_1(\omega))\Big].
\end{aligned}$$

Using (A.6) and the symmetry of T_k and U_k in $(\tilde{Y}, \tilde{Z})$ we get in particular that

$$\begin{aligned}
&\Big|\Big(\mathbb{P}^x \otimes \mathbb{P}^{x'}\Big)\Big(\tilde{Y}_t \in A, \tilde{Z}_t \in B\Big) - \Big(\mathbb{P}^x \otimes \mathbb{P}^{x'}\Big)\Big(\tilde{Y}_t \in B, \tilde{Z}_t \in A\Big)\Big| \\
&\leq 2\Big(\mathbb{P}^x \otimes \mathbb{P}^{x'}\Big)(t < U_n) \\
&\quad + \Big(\mathbb{P}^x \otimes \mathbb{P}^{x'}\Big)\Big[\mu^2_{(\tilde{Y}_{T_1(\omega)}, \tilde{Z}_{T_1(\omega)})}(dw, dw') \\
&\qquad \Big(\mathbb{P}^w \otimes \mathbb{P}^{w'}\Big)\Big(\tilde{Y}_{t-U_1(\omega)} \in [0,1]^{d-1}, \tilde{Z}_{t-U_1(\omega)} \in [0,1]^{d-1}\Big)\Big] \\
&\leq 2\Big(\mathbb{P}^x \otimes \mathbb{P}^{x'}\Big)(t < U_n) + \frac{3}{4},
\end{aligned}$$

the last by (A.7).

If $n \geq 2$ we can further condition the inner probability on $\mathcal{F}_{U_1}$ and decompose the last term in (A.10) into

$$\Big(\mathbb{P}^x \otimes \mathbb{P}^{x'}\Big)\Big[\mu^2_{\big(\tilde{Y}_{T_1(\omega)},\tilde{Z}_{T_1(\omega)}\big)}(dw,dw')\Big(\mathbb{P}^w \otimes \mathbb{P}^{w'}\Big)\Big(\tilde{Y}_{t-U_1(\omega)} \in A, \tilde{Z}_{t-U_1(\omega)} \in B,$$
$$t \geq U_1(\omega) + U_{n-1}\Big)1(t \geq U_1(\omega))\Big]$$
$$= \Big(\mathbb{P}^x \otimes \mathbb{P}^{x'}\Big)\Big[\mu^2_{\big(\tilde{Y}_{T_1(\omega)},\tilde{Z}_{T_1(\omega)}\big)}(dw,dw')\Big(\mathbb{P}^w \otimes \mathbb{P}^{w'}\Big)\Big[\mu^1_{\big(\tilde{Y}_{T_1(\omega')},\tilde{Z}_{T_1(\omega')}\big)}(dz,dz')$$
$$\Big(\mathbb{P}^z \otimes \mathbb{P}^{z'}\Big)\Big(\tilde{Y}_{t-(U_1(\omega)+U_1(\omega'))} \in A, \tilde{Z}_{t-(U_1(\omega)+U_1(\omega'))} \in B,$$
$$t \geq U_1(\omega) + U_1(\omega') + U_{n-2}\Big)1(t \geq U_1(\omega) + U_1(\omega'))\Big]1(t \geq U_1(\omega))\Big]$$
$$+ \Big(\mathbb{P}^x \otimes \mathbb{P}^{x'}\Big)\Big[\mu^2_{\big(\tilde{Y}_{T_1(\omega)},\tilde{Z}_{T_1(\omega)}\big)}(dw,dw')\Big(\mathbb{P}^w \otimes \mathbb{P}^{w'}\Big)\Big[\mu^2_{\big(\tilde{Y}_{T_1(\omega')},\tilde{Z}_{T_1(\omega')}\big)}(dz,dz')$$
$$\Big(\mathbb{P}^z \otimes \mathbb{P}^{z'}\Big)\Big(\tilde{Y}_{t-(U_1(\omega)+U_1(\omega'))} \in A, \tilde{Z}_{t-(U_1(\omega)+U_1(\omega'))} \in B,$$
$$t \geq U_1(\omega) + U_1(\omega') + U_{n-2}\Big)1(t \geq U_1(\omega) + U_1(\omega'))\Big]1(t \geq U_1(\omega))\Big].$$

Using this in (A.10) we obtain with (A.6)

$$\Big|\Big(\mathbb{P}^x \otimes \mathbb{P}^{x'}\Big)\Big(\tilde{Y}_t \in A, \tilde{Z}_t \in B\Big) - \Big(\mathbb{P}^x \otimes \mathbb{P}^{x'}\Big)\Big(\tilde{Y}_t \in B, \tilde{Z}_t \in A\Big)\Big|$$
$$\leq 2\Big(\mathbb{P}^x \otimes \mathbb{P}^{x'}\Big)(t < U_n)$$
$$+ \Big(\mathbb{P}^x \otimes \mathbb{P}^{x'}\Big)\Big[\mu^2_{\big(\tilde{Y}_{T_1(\omega)},\tilde{Z}_{T_1(\omega)}\big)}(dw,dw')\Big(\mathbb{P}^w \otimes \mathbb{P}^{w'}\Big)\Big[\mu^2_{\big(\tilde{Y}_{T_1(\omega')},\tilde{Z}_{T_1(\omega')}\big)}(dz,dz')$$
$$\Big(\mathbb{P}^z \otimes \mathbb{P}^{z'}\Big)\Big(\tilde{Y}_{t-(U_1(\omega)+U_1(\omega'))} \in [0,1]^{d-1}, \tilde{Z}_{t-(U_1(\omega)+U_1(\omega'))} \in [0,1]^{d-1}\Big)\Big]\Big]$$
$$\leq 2\Big(\mathbb{P}^x \otimes \mathbb{P}^{x'}\Big)(t < U_n) + \left(\frac{3}{4}\right)^2,$$

the last by (A.7).

By iterating the above decomposition we obtain for $n \in \mathbb{N}$ fixed,

$$\Big|\Big(\mathbb{P}^x \otimes \mathbb{P}^{x'}\Big)\Big(\tilde{Y}_t \in A, \tilde{Z}_t \in B\Big) - \Big(\mathbb{P}^x \otimes \mathbb{P}^{x'}\Big)\Big(\tilde{Y}_t \in B, \tilde{Z}_t \in A\Big)\Big|$$
$$\leq 2\Big(\mathbb{P}^x \otimes \mathbb{P}^{x'}\Big)(t < U_n) + \left(\frac{3}{4}\right)^n.$$

Recall that we have for almost all (x, x') w.r.t. $\tilde{\pi} \otimes \tilde{\pi}$ that almost surely $T_k < \infty$ for all $k \in \mathbb{N}$. Hence, to given $\epsilon > 0$ we can choose $n \in \mathbb{N}$ such that $\left(\frac{3}{4}\right)^n < \frac{\epsilon}{2}$ and then choose $T > 0$ such that $\Big(\mathbb{P}^x \otimes \mathbb{P}^{x'}\Big)(t < U_n) < \frac{\epsilon}{4}$ for all $t \geq T$. We obtain

$$\Big|\Big(\mathbb{P}^x \otimes \mathbb{P}^{x'}\Big)\Big(\tilde{Y}_t \in A, \tilde{Z}_t \in B\Big) - \Big(\mathbb{P}^x \otimes \mathbb{P}^{x'}\Big)\Big(\tilde{Y}_t \in B, \tilde{Z}_t \in A\Big)\Big| < \epsilon$$

for all $t \geq T$ and thus that

$$\lim_{t\to\infty} \sup_{A,B\in\mathcal{B}([0,1]^{d-1})} \Big| \Big(\mathbb{P}^x \otimes \mathbb{P}^{x'}\Big)\Big(\tilde{Y}_t \in A, \tilde{Z}_t \in B\Big) - \Big(\mathbb{P}^x \otimes \mathbb{P}^{x'}\Big)\Big(\tilde{Y}_t \in B, \tilde{Z}_t \in A\Big)\Big| = 0.$$

Choosing $A = [0,1]^{d-1}$ yields

$$\lim_{t\to\infty} \sup_{B\in\mathcal{B}([0,1]^{d-1})} \Big|\mathbb{P}^{x'}\Big(\tilde{Z}_t \in B\Big) - \mathbb{P}^x\Big(\tilde{Y}_t \in B\Big)\Big| = 0.$$

A simple tightness-argument completes the proof of our first step.

Next we shall extend our result $\mathcal{L}(\tilde{Y}_t|\tilde{Y}_0 = x) \Rightarrow \tilde{\pi}$ as $t \to \infty$ for all $x \in \tilde{S}$, $\tilde{\pi}$-a.s. to Lebesgue almost every $x \in \tilde{S}$. The proof goes by contradiction. Let $A = \{x \in \tilde{S} : \mathcal{L}(\tilde{Y}_t|\tilde{Y}_0 = x) \not\Rightarrow \tilde{\pi}\}$.

We first claim that A is Borel-measurable. Indeed, the martingale problem for $\tilde{Y}$ is well-posed and thus the process $\tilde{Y}$ is Feller continuous (see for example Stroock and Varadhan [4], Corollary 11.1.5). By using that the corresponding semigroup is a contraction we obtain the claim.

Suppose by contradiction that A has positive Lebesgue measure. In this case there exists a simply connected bounded open domain $D \subset \tilde{S}$ with smooth boundary such that $A \cap D$ has positive Lebesgue measure. As $\tilde{\pi}(A) = 0$ by the step above, $\tilde{\pi}(A \cap D) = 0$ follows. If $\tilde{Z}_t$ is the stationary solution of the SDE in $[0,1]^{d-1}$ started with initial law $\tilde{\pi}$, then $\mathbb{E}\Big[\int_0^T 1(\tilde{Z}_t \in A \cap D)dt\Big] = 0$ for all $T > 0$. On the other hand, by Theorem B.5 ("Occupation time measure for uniformly elliptic diffusions") of [1], we have for every $x \in D$, $\mathbb{E}\Big[\int_0^{\tau_D} 1(\tilde{Y}_t \in A \cap D)dt \mid \tilde{Y}_0 = x\Big] > 0$, where $\tau_D = \inf\{t \geq 0 : \tilde{Y}_t \notin D\}$. As $\tilde{\pi}$ assigns positive probability to D, we have

$$\int_D \mathbb{E}\Big[\int_0^{\tau_D} 1(\tilde{Y}_t \in A \cap D)dt \mid \tilde{Y}_0 = x\Big]\, \tilde{\pi}(dx) > 0.$$

By the monotone convergence theorem, we can choose T sufficiently large such that

$$\int_D \mathbb{E}\Bigg[\int_0^{\tau_D \wedge T} 1(\tilde{Y}_t \in A \cap D)dt \mid \tilde{Y}_0 = x\Bigg]\, \tilde{\pi}(dx) > 0.$$

But the l.h.s. is dominated by $\mathbb{E}\Big[\int_0^T 1(\tilde{Z}_t \in A \cap D)dt\Big] = 0$, which is a contradiction. Therefore A has Lebesgue measure zero.

It remains to show that $\mathcal{L}(\tilde{Y}_t|\tilde{Y}_0 = x) \Rightarrow \tilde{\pi}$ for all $x \in \tilde{S}$. Indeed, for $x \in \tilde{S}$, let $\epsilon > 0$ be such that $B_\epsilon(x) \subset \tilde{S}$. By Corollary B.3 applied to $D = B_\epsilon(x)$, the transition kernel $\mu_t^{B_\epsilon(x)}(x,\cdot)$ is absolutely continuous w.r.t. Lebesgue measure. As shown above, for Lebesgue almost every $y \in B_\epsilon(x)$, $\mathcal{L}(\tilde{Y}_{t+s}|\tilde{Y}_t = y) \Rightarrow \tilde{\pi}$ as $s \to \infty$. By observing that $\mu_t^{B_\epsilon(x)}(x, B_\epsilon(x)) \uparrow 1$ as $t \to 0$ (see (B.3)), we finally get $\mathcal{L}(\tilde{Y}_t|\tilde{Y}_0 = x) \Rightarrow \tilde{\pi}$ for arbitrary $x \in \tilde{S}$, which completes our proof. $\square$

Bibliography

[1] Dawson, D.A. and Greven, A. and den Hollander, F. and Sun, R. and Swart, J.M. The renormalization transformation for two-type branching models. *Ann. Inst. H. Poincaré Probab. Statist.* (2008) **44**, 1038–1077. MR2469334

[2] Ethier, S.N. and Kurtz, T.G. *Markov Processes: Characterization and Convergence*. Wiley and Sons, Inc., Hoboken, New Jersey , 2005. MR0838085

[3] Rogers, L.C.G. and Williams, D. *Diffusions, Markov Processes, and Martingales, vol. 2*, Reprint of the second (1994) edition. Cambridge Mathematical Univ. Press, Cambridge, 2000. MR1780932

[4] Stroock, D.W. and Varadhan, S.R.S. *Multidimensional Diffusion Processes*. Grundlehren Math. Wiss., vol. 233, Springer, Berlin-New York, 1979. MR532498

[5] Varadhan, S.R.S. *Probability Theory*. Courant Lect. Notes Math., 7, New York; Amer. Math. Soc., Providence, Rhode Island, 2001. MR1852999

Appendix B

Appendix for Chapter 4

The following Lemma and Corollary are necessary to prove Lemma 4.4.1 of Chapter 4.

Lemma B.0.1. *There exists* $N_0 < \infty$ *such that for all* $N \geq N_0, k \geq 1$

(a) $\left|\rho^k(t) - \exp\Big(-(1+o(1))\frac{kt^2}{6N}\Big)\right| \leq C\frac{1}{k}\exp\Big(-(1+o(1))\frac{kt^2}{12N}\Big)$ *for* $t \leq (1+o(1))\sqrt{\frac{N}{3}}$,

(b) $|\rho(t)| \leq \exp\Big(-C\frac{t^2}{12N}\Big)$ *for* $t \leq \left(\frac{6N}{(1+o(1))}\right)^{1/2}$,

(c) There exists $\delta > 0$ *such that* $|\rho(t)| \leq 1 - \delta$ *for* $t \in \left[\left(\frac{6N}{(1+o(1))}\right)^{1/2}, \pi N\right]$.

Proof. The proof mainly follows along the lines of the proof of Lemma 8 in Mueller and Tribe [3]. Some small changes ensued due to the different setup.

Recall the definition of $\rho(t)$ from equation (4.38).

For (b), we could not find the reference mentioned in [3] but the following reasoning in [3] based on applying Taylor's theorem at $t = 0$ works well without it.

For (a), first observe that $\rho^k(t) = \mathbb{E}\big[e^{itS_k}\big]$ and use Bhattacharya and Rao [1], (8.11), (8.13) and [1], Theorem 8.5. as suggested in [3]. We used that $\mathbb{E}[Y_1] = \mathbb{E}[Y_1^3] = 0$.

It remains to prove (c). We have to change the proof of [3], Lemma 8(c) slightly, as we used $x \not\sim x$. We get

$$\begin{aligned}
|\rho(t)| &= \left|\frac{1}{2c(N)N^{1/2}} \sum_{0<|j|\leq c(N)\sqrt{N}} e^{it\frac{j}{N}}\right| = \left|\frac{1}{2c(N)N^{1/2}} \sum_{0<j\leq c(N)\sqrt{N}} 2\mathrm{Re}\Big[e^{it\frac{j}{N}}\Big]\right| \\
&= \left|\frac{1}{c(N)N^{1/2}}\mathrm{Re}\left[\frac{e^{it\frac{1}{N}} - e^{it\frac{c(N)\sqrt{N}+1}{N}}}{1 - e^{it\frac{1}{N}}}\right]\right| \\
&= \left|\frac{1}{c(N)N^{1/2}}\mathrm{Re}\left[e^{-it\frac{1}{2N}}\frac{e^{it\frac{1}{N}} - e^{it\frac{c(N)\sqrt{N}+1}{N}}}{-2i\sin\big(\frac{t}{2N}\big)}\right]\right| \\
&\leq \frac{1}{c(N)N^{1/2}}\left|\frac{2}{2\sin\big(\frac{t}{2N}\big)}\right|.
\end{aligned}$$

For $\frac{1+\epsilon}{c(N)N^{1/2}} \le \frac{t}{2N} \le \frac{\pi}{2}$ with $\epsilon > 0$ fixed we get as an upper bound

$$\frac{1}{c(N)N^{1/2}} \left| \frac{1}{\sin\left(\frac{1+\epsilon}{c(N)N^{1/2}}\right)} \right| \le \frac{1}{1+\epsilon} < 1,$$

given N big enough. Finally use that $2 < \sqrt{6}$ to obtain the claim. □

Corollary B.0.2. *For $N \ge N_0$, $y \in N^{-1}\mathbb{Z}$ we have*

$$\left| N\mathbb{P}(S_k = y) - p\left((1+o(1))\frac{k}{3N}, y\right) \right| \le C_1 \left\{ N \exp\{-kC_2\} + N^{1/2}k^{-3/2} \right\},$$

where $C_1, C_2 > 0$ are some positive constants.

Proof. This result corresponds to Corollary 9 in [3]. The proof works similarly. Instead of the reference given at the beginning of the proof of Corollary 9 in [3], we used Durrett [2], p. 95, Ex. 3.2(ii) and [2], Thm. (3.3).

Note in particular that the result of Lemma B.0.1(c) can be extended to $t \in \left[(1+o(1))\sqrt{\frac{N}{3}}, \pi N\right]$ if we choose $\delta > 0$ small enough. Indeed, using Lemma B.0.1(b) we obtain

$$|\rho(t)| \le e^{-C\frac{t^2}{12N}} \le e^{-C\frac{N/3}{12N}} \le (1-\delta)$$

as claimed. □

Bibliography

[1] Bhattacharya, R.N. and Ranga Rao, R. *Normal approximation and asymptotic expansions.* Wiley and Sons, New York-London-Sydney, 1976. MR0436272

[2] Durrett, R. *Probability: Theory and Examples*, Third edition. Brooks/Cole-Thomsom Learning, Belmont, 2005.

[3] Mueller, C. and Tribe, R. Stochastic p.d.e.'s arising from the long range contact and long range voter processes. *Probab. Theory Related Fields* (1995) **102**, 519–545. MR1346264